

Second Edition

ADVANCED ENGINEERING MATHEMATICS

Erwin Kreyszig

PROFESSOR OF MATHEMATICS
OHIO STATE UNIVERSITY
COLUMBUS, OHIO



WILEY EASTERN PRIVATE LIMITED
NEW DELHI

First U.S. Edition, 1957
Second U.S. Edition, 1960

**Published by Anand R. Kundaji for Wiley Eastern Private Limited,
J 41 South Extension 1, New Delhi 49, and printed by G.D. Makhija
at India Offset Press, A-I, Maya Puri Industrial Estate, New Delhi-27
Printed in India.**

PREFACE

Purpose of the book. This book is intended to introduce students of engineering and physics to those fields of mathematics which, from the modern point of view, seem to be the most important in connection with practical problems. Topics are chosen according to the frequency of occurrence in applications. New ideas of modern mathematical training, as expressed in various recent symposia on engineering education, were taken into account. The book should suit those institutions that have offered extended mathematical training for a long time as well as those that intend to follow the general trend of broadening the program of instruction in mathematics.

A course in elementary calculus is the sole prerequisite.

The material included in the book has formed the basis of various lecture courses given to undergraduate and graduate students of engineering, physics, and mathematics in this country, in Canada, and in Europe.

Changes in the second edition. The present edition differs essentially from the first edition as follows.

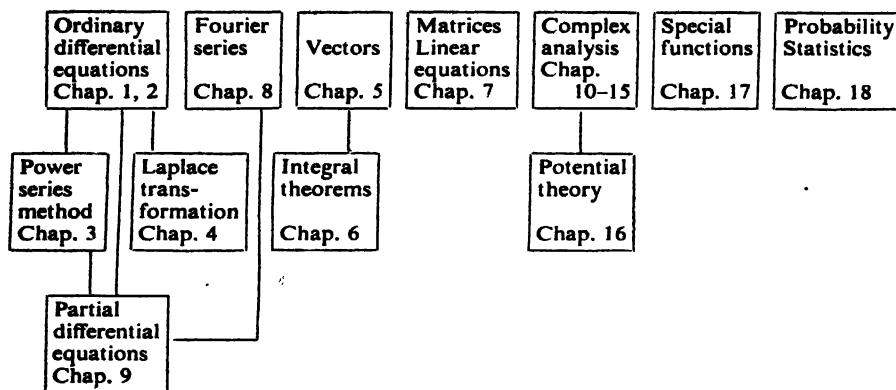
The problems have been changed.

A new chapter on probability and engineering statistics has been added. It is suitable for a one-semester course meeting 3 or 4 hours a week and is completely independent of the other chapters in the book.

The arrangement of the material in complex analysis has been changed to make the presentation more simple and compact. The consideration of complex series in general is now immediately followed by that on Taylor and Laurent series without interruption by other topics.

Simplifications and improvements have been made at various other places, in particular in Chaps. 3, 4, and 7.

Content and arrangement. The arrangement of the subject matter in major parts can be seen from the diagram on the next page.



Much space is devoted to ordinary differential equations and complex analysis, probably the two most important fields for engineers. But also the length of the other chapters on vectors, Fourier series, matrices, statistics, etc., is such that the representation can be used as a text for courses of the usual type.

Sections which may be omitted in a shorter course are indicated at the beginning of each chapter.

To facilitate the use of parts of the book, the chapters are kept as independent of one another as possible.

The chapters are subdivided into relatively short sections. Each section includes typical examples and problems illustrating the concepts, methods, and results and their engineering applications.

Historical notes, references to original literature, and more than 400 figures are included in the text.

Problems and answers. The book contains more than 3500 carefully selected problems, which range from simple routine exercises to practical applications of considerable complexity. Answers to odd-numbered problems are included at the end of the book, starting on page 845.

References. A list of some books for reference and further study can be found at the end of the book, on page 841.

Suggestions for courses. The material may be taken in sequence and is then suitable for four consecutive semester courses, meeting 3–5 hours a week as follows:

First semester. Ordinary differential equations (Chaps. 1–4).

Second semester. Vectors and matrices (Chaps. 5–7).

Third semester. Fourier series and partial differential equations (Chaps. 8, 9).

Fourth semester. Complex analysis (Chaps. 10–16).

A course on ENGINEERING STATISTICS (3–5 hours a week; Chap. 18) may be taught during any of those four semesters or afterwards.

The book is also suitable for various independent one-semester courses meeting 3 hours a week; for example:

- Introduction to ordinary differential equations (Chaps. 1, 2).
- Laplace transformation (Chap. 4)
- Vector calculus (Chaps. 5, 6).
- Matrices and systems of linear equations (Chap. 7).
- Fourier series and integrals (Chap. 8).
- Partial differential equations (Chap. 9).
- Complex analysis (Chaps. 10–15).
- Infinite series (Secs. 13.1–13.5, 8.1–8.6, 17.3, 17.4).
- Special functions (Chap. 17, Secs. 3.3–3.7, 8.12, 8.13).
- Numerical analysis (Secs. 0.6–0.8, 1.13, 2.17, 7.14, 7.15, 8.8, 17.3, 17.4).

Principles for selection of topics. Which topics should be contained in a book of the present type, and how should these topics be arranged and presented?

To find some answer to these basic questions we may take a look at the historical development of engineering mathematics. This development shows the following two interesting features:

1. Mathematics has become more and more important in engineering science, and it is easy to conjecture that this trend will also continue in the future. An important reason for this tendency is that the problems in modern engineering are so complex that most of them cannot be solved solely on the basis of physical intuition and past experience. This empirical approach, which has been successful for many problems, fails as soon as high speeds, large forces, high temperatures or other abnormal conditions are involved, and the situation becomes still more critical by the fact that various modern materials (plastics, alloys, etc.) have unusual physical properties. Experimental work becomes complicated, time-consuming, and expensive. Here mathematics offers help in planning constructions and experiments, and in evaluating experimental data.

2. Mathematical methods which were developed for purely theoretical reasons suddenly became of great importance in engineering mathematics. Examples are the theory of matrices, conformal mapping, and the theory of differential equations having periodic solutions.

What are the reflections of this situation on the teaching of engineering mathematics? Since there will be more and more mathematics needed, should we try to include more and more topics in our courses, devoting less and less time to each topic? Or should we concentrate on a few carefully selected basic things of general practical importance which are suitable for teaching the student mathematical thinking and developing his own creative ability?

Sixty or eighty years ago no one was able to predict that conformal mapping or matrices would ever be of importance in the mathematical part of engineering work. Similarly, it is difficult to conjecture which new mathematical theories will have applications to engineering twenty or thirty years from now. But no matter what happens in that respect, if a student has a good training in the fundamentals of mathematics he will meet the future needs because he will be able to get acquainted with new methods by his own further study.

It follows that the most important objective and purpose in engineering mathematics seems to be that the student becomes familiar with mathematical thinking. He should learn to recognize the guiding principles and ideas "behind the scenes," which are more important than formal manipulations. He should get the impression that mathematics is not a collection of tricks and recipes but a systematic science of practical importance, resting on a relatively small number of basic concepts and involving powerful unifying methods. He should soon convince himself of the necessity for applying mathematical procedures to engineering problems, and he will find that the theory and its applications are related to each other like a tree and its fruits.

The student will see that the application of mathematics to an engineering problem consists essentially of three phases:

1. *Translation of the given physical information into a mathematical form. In this way we obtain a mathematical model of the physical situation. This model may be a differential equation, a system of linear equations, or some other mathematical expression.*

2. *Treatment of the model by mathematical methods. This will lead to the solution of the given problem in mathematical form.*

3. *Interpretation of the mathematical result in physical terms.*

All three steps seem to be of equal importance, and the presentation in this book is such that it will help the student to develop skill in carrying out all these steps. In this connection, preference has been given to applications which are of a general nature.

In some considerations it will be unavoidable to rely on results whose proofs are beyond the level of a book of the present type. In any case these points are marked distinctly, because hiding difficulties and oversimplifying matters is no real help to the student in preparing him for his professional work.

These are some of the guiding principles I used in selecting and presenting the material included in this book. The choice has been made with greatest care, using past and present teaching and research experience and resisting the temptation to consider "everything which is important" in engineering mathematics.

Particular efforts have been made in presenting the topics as simply, clearly, and accurately as possible; this refers also to the choice of the

notations. In each chapter the level increases gradually, avoiding jumps and accumulations of difficult theoretical considerations.

Numerical methods. Engineering mathematics comes down ultimately to numerical results, and the engineering student should therefore supplement his mathematical equipment with a definite knowledge of some fundamental numerical procedures. For this reason, numerical work is included in the text as well as in the problems, and basic numerical methods are considered in connection with differential equations, matrices, Fourier series, etc. Preference is given to methods suitable for numerical work in a simple form (slide rule, desk calculator) as well as in connection with modern high-speed computation.

Acknowledgment. I am indebted to many of my former teachers, colleagues, and students for advice and help in preparing this book. Various parts of the manuscript were distributed to my classes in mimeographed form and returned to me, together with suggestions for improvement. Discussions with various engineers and mathematicians were of great help to me; I want to mention in particular Professors S. Bergman, P. L. Chambré, A. Cronheim, J. W. Dettman, R. G. Helsel, E. C. Klipple, H. Kuhn, H. B. Mann, I. Marx, W. D. Munroe, H.-W. Pu, T. Rado, H. A. Smith, J. P. Spencer, J. Todd, H. J. Weiss, A. Wilansky, all in this country, Professor H. S. M. Coxeter of Toronto, and Professors B. Baule, H. Behnke, H. Florian, H. Graf, F. Hohenberg, K. Klotter, M. Pinl, F. Reutter, C. Schmieden, H. Unger, A. Walther, H. Wielandt, all in Europe. I can offer here only an inadequate acknowledgment of my appreciation.

Last, but not least, I want to thank John Wiley and Sons for their effective cooperation and their great care in preparing this edition of the book.

Suggestions of many readers were evaluated in preparing the present edition. Any further comment and suggestion for improvement of the book will be gratefully received.

Erwin Kreyszig

CONTENTS

Introduction	<i>Review of Some Topics From Algebra and Calculus</i>	1
0.1	Elementary Functions, 1	
0.2	Partial Derivatives, 8	
0.3	Second- and Third-Order Determinants, 10	
0.4	Complex Numbers, 18	
0.5	Polar Form of Complex Numbers, 22	
0.6	Some General Remarks About Numerical Computations, 24	
0.7	Solution of Equations, 26	
0.8	Approximate Integration, 31	
Chapter 1	<i>Ordinary Differential Equations of the First Order</i>	39
1.1	Basic Concepts and Ideas, 39	
1.2	Geometrical Considerations. Isoclines, 46	
1.3	Separable Equations, 49	
1.4	Equations Reducible to Separable Form, 57	
1.5	Exact Differential Equations, 59	
1.6	Integrating Factors, 61	
1.7	Linear First-Order Differential Equations, 63	
1.8	Variation of Parameters, 67	
1.9	Electric Circuits, 69	
1.10	Families of Curves. Orthogonal Trajectories, 74	
1.11	Picard's Iteration Method, 79	
1.12	Existence and Uniqueness of Solutions, 82	
1.13	Numerical Methods for Differential Equations of the First Order, 86	

Chapter 2 Ordinary Linear Differential Equations	93
2.1 Homogeneous Linear Equations of the Second Order, 94	
2.2 Homogeneous Second-Order Equations with Constant Coefficients, 97	
2.3 General Solution. Fundamental System, 99	
2.4 Complex Roots of the Characteristic Equation. Initial Value Problem, 102	
2.5 Double Root of the Characteristic Equation, 106	
2.6 Free Oscillations, 108	
2.7 Cauchy Equation, 116	
2.8 Existence and Uniqueness of Solutions, 117	
2.9 Homogeneous Linear Equations of Arbitrary Order, 123	
2.10 Homogeneous Linear Equations of Arbitrary Order with Constant Coefficients, 126	
2.11 Nonhomogeneous Linear Equations, 128	
2.12 A Method for Solving Nonhomogeneous Linear Equations, 130	
2.13 Forced Oscillations. Resonance, 134	
2.14 Electric Circuits, 140	
2.15 Complex Method for Obtaining Particular Solutions, 144	
2.16 General Method for Solving Nonhomogeneous Equations, 147	
2.17 Numerical Methods for Second-Order Differential Equations, 149	
Chapter 3 Power Series Solutions of Differential Equations	155
3.1 The Power Series Method, 155	
3.2 Theoretical Basis of the Power Series Method, 159	
3.3 Legendre's Equation Legendre Polynomials, 164	
3.4 Extended Power Series Method Indicial Equation, 168	
3.5 Bessel's Equation Bessel Functions of the First Kind, 179	
3.6 Further Properties of Bessel Functions of the First Kind. 185	
3.7 Bessel Functions of the Second Kind, 187	
Chapter 4 Laplace Transformation	192
4.1 Laplace Transform. Inverse Transform. Linearity, 192	
4.2 Laplace Transforms of Derivatives and Integrals, 198	
4.3 Transformation of Ordinary Differential Equations, 201	
4.4 Partial Fractions, 204	
4.5 Examples and Applications, 209	
4.6 Differentiation and Integration of Transforms, 213	
4.7 Unit Step Function, 216	
4.8 Shifting on the t -axis, 219	
4.9 Periodic Functions, 223	
4.10 Table 17. Some Laplace Transforms, 231	

Chapter 5 Vector Analysis **235**

- 5.1 Scalars and Vectors, 235**
- 5.2 Components of a Vector, 237**
- 5.3 Vector Addition. Multiplication by Scalars, 239**
- 5.4 Scalar Product, 242**
- 5.5 Vector Product, 247**
- 5.6 Vector Products in Terms of Components, 248**
- 5.7 Scalar Triple Product. Linear Dependence of Vectors, 254**
- 5.8 Other Repeated Products, 258**
- 5.9 Scalar Fields and Vector Fields, 260**
- 5.10 Vector Calculus, 263**
- 5.11 Curves, 265**
- 5.12 Arc Length, 268**
- 5.13 Tangent. Curvature and Torsion, 270**
- 5.14 Velocity and Acceleration, 274**
- 5.15 Chain Rule and Mean Value Theorem for Functions of Several Variables, 278**
- 5.16 Directional Derivative. Gradient of a Scalar Field, 282**
- 5.17 Transformation of Coordinate Systems and Vector Components, 288**
- 5.18 Divergence of a Vector Field, 292**
- 5.19 Curl of a Vector Field, 296**

Chapter 6 Line and Surface Integrals. Integral Theorems **299**

- 6.1 Line Integral, 299**
- 6.2 Evaluation of Line Integrals, 302**
- 6.3 Double Integrals, 306**
- 6.4 Transformation of Double Integrals into Line Integrals, 313**
- 6.5 Surfaces, 318**
- 6.6 Tangent Plane First Fundamental Form. Area, 321**
- 6.7 Surface Integrals, 327**
- 6.8 Triple Integrals. Divergence Theorem of Gauss, 332**
- 6.9 Consequences and Applications of the Divergence Theorem, 336**
- 6.10 Stokes's Theorem, 342**
- 6.11 Consequences and Applications of Stokes's Theorem, 346**
- 6.12 Line Integrals Independent of Path, 347**

Chapter 7 Matrices and Determinants.
Systems of Linear Equations **356**

- 7.1 Basic Concepts. Addition of Matrices, 356**
- 7.2 Matrix Multiplication, 361**

- 7.3 Determinants, 368
- 7.4 Submatrices. Rank, 380
- 7.5 Systems of n Linear Equations in n Unknowns. Cramer's Rule, 382
- 7.6 Arbitrary Homogeneous Systems of Linear Equations, 386
- 7.7 Arbitrary Nonhomogeneous Systems of Linear Equations, 392
- 7.8 Further Properties of Systems of Linear Equations, 395
- 7.9 Gauss's Elimination Method, 398
- 7.10 The Inverse of a Matrix, 401
- 7.11 Eigenvalues. Eigenvectors, 406
- 7.12 Bilinear, Quadratic, Hermitian, and Skew-Hermitian Forms, 413
- 7.13 Eigenvalues of Hermitian, Skew-Hermitian, and Unitary Matrices, 417
- 7.14 Bounds for Eigenvalues, 422
- 7.15 Determination of Eigenvalues by Iteration, 426

Chapter 8 *Fourier Series and Integrals*

431

- 8.1 Periodic Functions. Trigonometric Series, 431
- 8.2 Fourier Series. Euler's Formulas, 434
- 8.3 Even and Odd Functions, 440
- 8.4 Functions Having Arbitrary Period, 444
- 8.5 Half-Range Expansions, 447
- 8.6 Determination of Fourier Coefficients Without Integration, 451
- 8.7 Forced Oscillations, 455
- 8.8 Numerical Methods for Determining Fourier Coefficients. Square Error, 458
- 8.9 Instrumental Methods for Determining Fourier Coefficients, 464
- 8.10 The Fourier Integral, 465
- 8.11 Orthogonal Functions, 473
- 8.12 Sturm-Liouville Problem, 476
- 8.13 Orthogonality of Bessel Functions, 483

Chapter 9 *Partial Differential Equations*

486

- 9.1 Basic Concepts, 486
- 9.2 Vibrating String. One-Dimensional Wave Equation, 488
- 9.3 Separation of Variables (Product Method), 490
- 9.4 D'Alembert's Solution of the Wave Equation, 499
- 9.5 One-Dimensional Heat Flow, 501
- 9.6 Heat Flow in an Infinite Bar, 506
- 9.7 Vibrating Membrane. Two-Dimensional Wave Equation, 510
- 9.8 Rectangular Membrane, 512
- 9.9 Laplacian in Polar Coordinates, 519
- 9.10 Circular Membrane. Bessel's Equation, 521

9.11 Laplace's Equation. Potential, 526	
9.12 Laplace's Equation in Spherical Coordinates. Legendre's Equation, 529	
Chapter 10 <i>Complex Analytic Functions</i>	534
10.1 Complex Numbers. Triangle Inequality, 535	
10.2 Limit. Derivative. Analytic Function, 539	
10.3 Cauchy-Riemann Equations. Laplace's Equation, 543	
10.4 Rational Functions. Root, 547	
10.5 Exponential Function, 550	
10.6 Trigonometric and Hyperbolic Functions, 553	
10.7 Logarithm. General Power, 556	
-Chapter 11 <i>Conformal Mapping</i>	560
11.1 Mapping, 560	
11.2 Conformal Mapping, 564	
11.3 Linear Transformations, 568	
11.4 Special Linear Transformations, 572	
11.5 Mapping by Other Elementary Functions, 577	
11.6 Riemann Surfaces, 584	
Chapter 12 <i>Complex Integrals</i>	588
12.1 Line Integral in the Complex Plane, 588	
12.2 Basic Properties of the Complex Line Integral, 594	
12.3 Cauchy's Integral Theorem, 595	
12.4 Evaluation of Line Integrals by Indefinite Integration, 602	
12.5 Cauchy's Integral Formula, 605	
12.6 The Derivatives of an Analytic Function, 607	
Chapter 13 <i>Sequences and Series</i>	611
13.1 Sequences, 611	
13.2 Series, 618	
13.3 Tests for Convergence and Divergence of Series, 623	
13.4 Operations on Series, 629	
13.5 Power Series, 633	
13.6 Functions Represented by Power Series, 640	

Chapter 14	<i>Taylor and Laurent Series</i>	646
14.1	Taylor Series, 646	
14.2	Taylor Series of Elementary Functions, 650	
14.3	Practical Methods for Obtaining Power Series, 652	
14.4	Uniform Convergence, 655	
14.5	Laurent Series, 663	
14.6	Behavior of Functions at Infinity, 668	
Chapter 15	<i>Integration by the Method of Residues</i>	671
15.1	Zeros and Singularities, 671	
15.2	Residues, 675	
15.3	The Residue Theorem, 678	
15.4	Evaluation of Real Integrals, 680	
Chapter 16	<i>Complex Analytic Functions and Potential Theory</i>	689
16.1	Electrostatic Fields, 689	
16.2	Two-Dimensional Fluid Flow, 693	
16.3	Special Complex Potentials, 697	
16.4	General Properties of Harmonic Functions, 702	
16.5	Poisson's Integral Formula, 705	
Chapter 17	<i>Special Functions. Asymptotic Expansions</i>	
17.1	Gamma and Beta Functions, 710	
17.2	Error Function Fresnel Integrals. Sine and Cosine Integrals, 716	
17.3	Asymptotic Expansions, 720	
17.4	Further Properties of Asymptotic Expansions, 725	
Chapter 18	<i>Probability and Statistics</i>	
18.1	Nature and Purpose of Mathematical Statistics, 732	
18.2	Tabular and Graphical Representation of Samples, 734	
18.3	Sample Mean and Sample Variance, 740	
18.4	Random Experiments, Outcomes, Events, 744	
18.5	Probability, 748	
18.6	Permutations and Combinations, 752	
18.7	Random Variables. Discrete and Continuous Distributions, 756	
18.8	Mean and Variance of a Distribution, 762	
18.9	Binomial, Poisson, and Hypergeometric Distributions, 766	
18.10	Normal Distribution, 770	

18.11 Distributions of Several Random Variables, 777	
18.12 Random Sampling. Random Numbers, 783	
18.13 Estimation of Parameters, 785	
18.14 Confidence Intervals, 790	
18.15 Testing of Hypotheses, Decisions, 799	
18.16 Quality Control, 808	
18.17 Acceptance Sampling, 811	
18.18 Goodness of Fit. χ^2 -Test, 817	
18.19 Nonparametric Tests, 819	
18.20 Pairs of Measurements. Fitting Straight Lines, 822	
18.21 Statistical Tables, 828	
Appendix 1 <i>References</i>	841
Appendix 2 <i>Answers to Odd-Numbered Problems</i>	845
Index	882

Advanced Engineering Mathematics

INTRODUCTION

Review of Some Topics from Algebra and Calculus

This introductory chapter includes some topics which are usually covered in elementary algebra and calculus. We shall refer to sections of this chapter whenever we need some of these topics as a prerequisite for our further consideration.

References: Appendix 1, Ref. [A13].

Answers to problems: Appendix 2.

0.1 Elementary Functions

In this section we shall present a collection of some basic formulas for reference.

Figure 1 shows the graph of the exponential function e^x , where

$$e = 2.7\ 1828\ 1828\ 45\ 90\ 45\ 2 \dots$$

Basic identities are

$$(1) \quad e^x e^y = e^{x+y}, \quad e^x/e^y = e^{x-y}, \\ (e^x)^y = e^{xy}.$$

The inverse of e^x is the natural logarithm $\ln x$ (Fig. 2). It satisfies the identities

$$(2) \quad \ln(xy) = \ln x + \ln y, \quad \ln \frac{x}{y} = \ln x - \ln y, \\ \ln(x^a) = a \ln x.$$

Furthermore,

$$e^{\ln x} = x, \quad e^{-\ln x} = e^{\ln(1/x)} = \frac{1}{x}.$$

The inverse of the exponential function 10^x is the *logarithm of base 10* which is denoted by

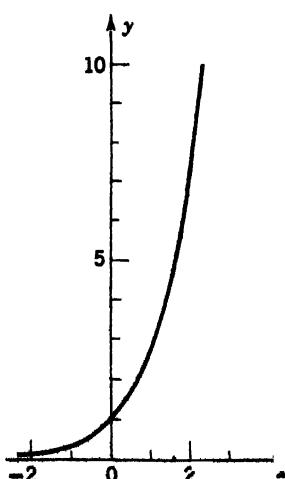
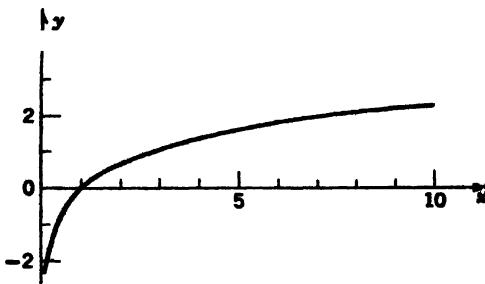


Fig. 1. Exponential function e^x .

Fig. 2. Natural logarithm $\ln x$.

$\log_{10} x$ or simply by $\log x$. We have

$$\log x = M \ln x$$

where

$$M = \log e = 0.43429\ 44819\ 03251\ 82765$$

and conversely

$$\ln x = \frac{1}{M} \log x$$

where

$$\frac{1}{M} = \ln 10 = 2.30258\ 50929\ 94045\ 68402.$$

The sine and cosine functions $\sin x$ and $\cos x$ are defined in trigonometry for all values of x . Throughout calculus, angles are measured in radians so that both functions have the period 2π .

A function $w = f(x)$ which is defined for all x and has the property

$$(3) \quad f(-x) = f(x) \quad \text{for all } x$$

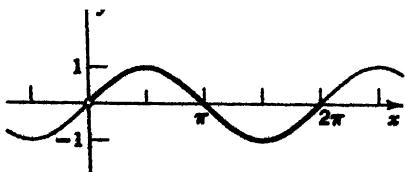
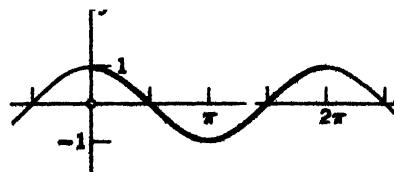
is called an even function. The graph of such a function is symmetric with respect to the w -axis. If $f(x)$ is defined for all x and

$$(4) \quad f(-x) = -f(x) \quad \text{for all } x,$$

then $f(x)$ is called an odd function. These are two quite important concepts.

Since

$$(5) \quad \sin(-x) = -\sin x, \quad \cos(-x) = \cos x$$

Fig. 3. $\sin x$.Fig. 4. $\cos x$.

Section 0.1]

Elementary Functions

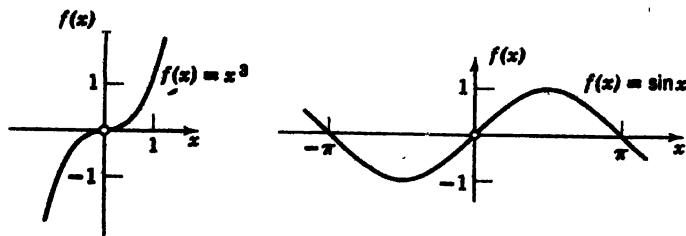


Fig. 5. Odd functions.

$\sin x$ is odd while $\cos x$ is even. The exponential function e^x is neither odd nor even.

The functions $\sin x$ and $\cos x$ are related by the identity

$$(6) \quad \sin^2 x + \cos^2 x = 1.$$

The addition formulas of the sine function are

$$(7) \quad \begin{aligned} \sin(x+y) &= \sin x \cos y + \cos x \sin y \\ \sin(x-y) &= \sin x \cos y - \cos x \sin y. \end{aligned}$$

In particular,

$$(7^*) \quad \sin 2x = 2 \sin x \cos x.$$

The addition formulas of the cosine function are

$$(8) \quad \begin{aligned} \cos(x+y) &= \cos x \cos y - \sin x \sin y \\ \cos(x-y) &= \cos x \cos y + \sin x \sin y. \end{aligned}$$

In particular,

$$(8^*) \quad \cos 2x = \cos^2 x - \sin^2 x.$$

From (7) and (8),

$$(9) \quad \begin{aligned} \sin x &= \cos\left(x - \frac{\pi}{2}\right) = \cos\left(\frac{\pi}{2} - x\right) \\ \cos x &= \sin\left(x + \frac{\pi}{2}\right) = \sin\left(\frac{\pi}{2} - x\right) \\ \sin(\pi - x) &= \sin x, \quad \cos(\pi - x) = -\cos x. \end{aligned}$$

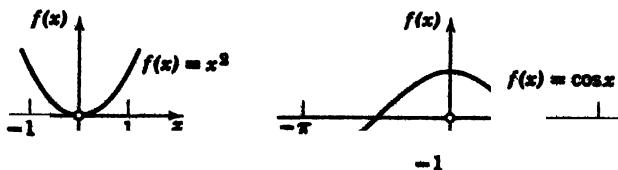


Fig. 6. Even functions.

From (8*) and (6),

$$(10) \quad \cos^2 x = \frac{1}{2}(1 + \cos 2x), \quad \sin^2 x = \frac{1}{2}(1 - \cos 2x).$$

From (8),

$$(11a) \quad \begin{aligned} \sin x \sin y &= \frac{1}{2}[-\cos(x+y) + \cos(x-y)] \\ \cos x \cos y &= \frac{1}{2}[\cos(x+y) + \cos(x-y)], \end{aligned}$$

and from (7),

$$(11b) \quad \sin x \cos y = \frac{1}{2}[\sin(x+y) + \sin(x-y)].$$

From this, by setting $x+y=u$ and $x-y=v$,

$$\begin{aligned} \sin u + \sin v &= 2 \sin \frac{u+v}{2} \cos \frac{u-v}{2} \\ (12) \quad \cos u + \cos v &= 2 \cos \frac{u+v}{2} \cos \frac{u-v}{2} \\ \cos v - \cos u &= 2 \sin \frac{u+v}{2} \sin \frac{u-v}{2}. \end{aligned}$$

The other trigonometric functions are defined by the identities

$$(13) \quad \tan x = \frac{\sin x}{\cos x}, \quad \cot x = \frac{\cos x}{\sin x}, \quad \sec x = \frac{1}{\cos x}, \quad \csc x = \frac{1}{\sin x}$$

The addition formulas for the tangent are

$$(14) \quad \tan(x+y) = \frac{\tan x + \tan y}{1 - \tan x \tan y}, \quad \tan(x-y) = \frac{\tan x - \tan y}{1 + \tan x \tan y}$$

In applications it is sometimes required to write $A \cos x + B \sin x$ where A and B are given constants, in the form $C \cos(x-\delta)$ where C and δ are

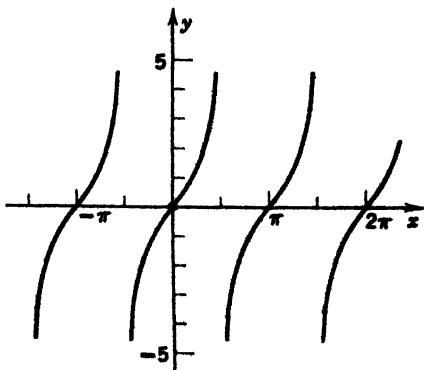


Fig. 7. $\tan x$.

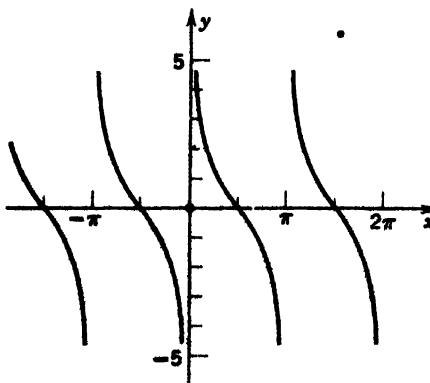
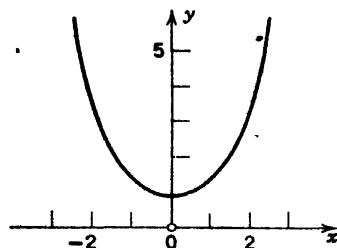
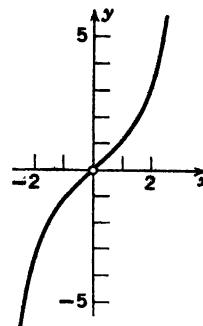


Fig. 8. $\cot x$.

Fig. 9. $\cosh x$.Fig. 10. $\sinh x$.

constants. From (8) we obtain

$$C \cos(x - \delta) = C \cos \delta \cos x + C \sin \delta \sin x$$

and this is equal to $A \cos x + B \sin x$, if $C \cos \delta = A$ and $C \sin \delta = B$. Using (6), we thus obtain

$$(15a) \quad A \cos x + B \sin x = \sqrt{A^2 + B^2} \cos(x - \delta)$$

where

$$\tan \delta = \frac{\sin \delta}{\cos \delta} = \frac{B}{A}.$$

Similarly,

$$(15b) \quad A \cos x + B \sin x = \sqrt{A^2 + B^2} \sin(x \pm \delta)$$

where

$$\tan \delta = \frac{\sin \delta}{\cos \delta} = \pm \frac{A}{B}.$$

The hyperbolic cosine and sine functions are defined by the identities

$$(16) \quad \cosh x = \frac{1}{2}(e^x + e^{-x}), \quad \sinh x = \frac{1}{2}(e^x - e^{-x}).$$

$\cosh x$ is even, while $\sinh x$ is odd. The other hyperbolic functions are defined by the identities

$$(17) \quad \tanh x = \frac{\sinh x}{\cosh x}, \quad \coth x = \frac{\cosh x}{\sinh x}.$$

From the definitions we obtain

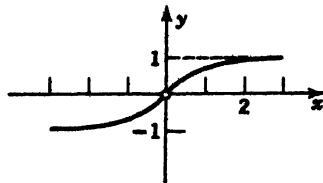
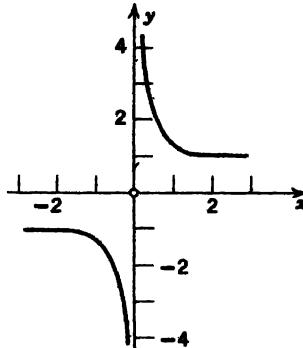
$$\cosh^2 x - \sinh^2 x = 1$$

$$(18) \quad \cosh x + \sinh x = e^x, \quad \cosh x - \sinh x = e^{-x}$$

Table 1. Some Elementary Functions

x	$\sin x$	$\cos x$	$\tan x$	e^x	$\sinh x$	$\cosh x$
0	0.00000	1.00000	0.00000	1.00000	0.00000	1.00000
0.1	0.09983	0.99500	0.10033	1.10517	0.10017	1.00500
0.2	0.19867	0.98007	0.20271	1.22140	0.20134	1.02007
0.3	0.29552	0.95534	0.30934	1.34986	0.30452	1.04534
0.4	0.38942	0.92106	0.42279	1.49182	0.41075	1.08107
0.5	0.47943	0.87758	0.54630	1.64872	0.52110	1.12763
0.6	0.56464	0.82534	0.68414	1.82212	0.63665	1.18547
0.7	0.64422	0.76484	0.84229	2.01375	0.75858	1.25517
0.8	0.71736	0.69671	1.02964	2.22554	0.88811	1.33743
0.9	0.78333	0.62161	1.26016	2.45960	1.02652	1.43309
1.0	0.84147	0.54030	1.55741	2.71828	1.17520	1.54308
1.1	0.89121	0.45360	1.96476	3.00417	1.33565	1.66852
1.2	0.93204	0.36236	2.57215	3.32012	1.50946	1.81066
1.3	0.96356	0.26750	3.60210	3.66930	1.69838	1.97091
1.4	0.98545	0.16997	5.79788	4.05520	1.90430	2.15090
1.5	0.99750	0.07074	14.10142	4.48169	2.12928	2.35241
1.6	0.99957	-0.02920	-34.23253	4.95303	2.37557	2.57746
1.7	0.99166	-0.12884	-7.69660	5.47395	2.64563	2.82832
1.8	0.97385	-0.22720	-4.28626	6.04965	2.94217	3.10747
1.9	0.94630	-0.32329	-2.92710	6.68589	3.26816	3.41773
2.0	0.90930	-0.41615	-2.18504	7.38906	3.62686	3.76220

x	$\ln x$						
1.0	0.00000	2.0	0.69315	3.0	1.09861	5	1.60944
1.1	0.09531	2.1	0.74194	3.1	1.13140	7	1.94591
1.2	0.18232	2.2	0.78846	3.2	1.16315	11	2.39790
1.3	0.26236	2.3	0.83291	3.3	1.19392	13	2.56495
1.4	0.33647	2.4	0.87547	3.4	1.22378	17	2.83321
1.5	0.40547	2.5	0.91629	3.5	1.25276	19	2.94444
1.6	0.47000	2.6	0.95551	3.6	1.28093	23	3.13549
1.7	0.53063	2.7	0.99325	3.7	1.30833	29	3.36730
1.8	0.58779	2.8	1.02962	3.8	1.33500	31	3.43399
1.9	0.64185	2.9	1.06471	3.9	1.36098	37	3.61092

Fig. 11. $\tanh x$.Fig. 12. $\coth x$.

and furthermore the addition formulas

$$(19) \quad \begin{aligned} \sinh(x \pm y) &= \sinh x \cosh y \pm \cosh x \sinh y \\ \cosh(x \pm y) &= \cosh x \cosh y \pm \sinh x \sinh y \\ \tanh(x \pm y) &= \frac{\tanh x \pm \tanh y}{1 \pm \tanh x \tanh y}. \end{aligned}$$

Problems

Are the following functions even, odd, or neither even nor odd?

1. $\sin(x^2)$ 2. $\sin^2 x$ 3. $\sin x + \cos x$ 4. x, x^3, x^5

5. $\tan x$ 6. $\tanh x$ 7. x^2, x^4, x^6 8. $\ln(1 + e^x) - \frac{x}{2}$

9. Prove that the sum and the product of even functions are even functions.

10. Prove that the sum of odd functions is odd and the product of two odd functions is even.

11. Derive (9) from (7) and (8).

12. Derive (11) from (7) and (8).

13. Prove (18).

14. Prove the addition formulas of the hyperbolic sine and cosine.

Prove the following identities.

15. $\sinh x \sinh y = \frac{1}{2}[\cosh(x + y) - \cosh(x - y)]$

16. $\cosh x \cosh y = \frac{1}{2}[\cosh(x + y) + \cosh(x - y)]$

17. $\sinh x \cosh y = \frac{1}{2}[\sinh(x + y) + \sinh(x - y)]$

18. $\sinh^2 x = \frac{1}{2}(\cosh 2x - 1), \cosh^2 x = \frac{1}{2}(\cosh 2x + 1)$

19. Using the differentiation formula of e^x , find the derivatives of $\sinh x$, $\cosh x$, and $\tanh x$.

20. Show that for large x , $\sinh x \approx e^x/2$, $\cosh x \approx e^x/2$.

0.2 Partial Derivatives

Let $z = f(x, y)$ be a real function of two independent real variables, x and y . If we hold y constant, say, $y = y_1$, and think of x as a variable, then $f(x, y_1)$ depends on x alone. If the derivative of $f(x, y_1)$ with respect to x for a value $x = x_1$ exists, then the value of this derivative is called the *partial derivative of $f(x, y)$ with respect to x at the point (x_1, y_1)* and is denoted by

$$\frac{\partial f}{\partial x} \Big|_{(x_1, y_1)} \quad \text{or by} \quad \frac{\partial z}{\partial x} \Big|_{(x_1, y_1)}$$

Other notations are

$$f_x(x_1, y_1) \quad \text{and} \quad z_x(x_1, y_1);$$

these may be used when subscripts are not used for another purpose and there is no danger of confusion.

We thus have, by the definition of the derivative,

$$(1) \quad \frac{\partial f}{\partial x} \Big|_{(x_1, y_1)} = \lim_{\Delta x \rightarrow 0} \frac{f(x_1 + \Delta x, y_1) - f(x_1, y_1)}{\Delta x}$$

The partial derivative of $z = f(x, y)$ with respect to y is defined similarly; we now hold x constant, say, equal to x_1 , and differentiate $f(x_1, y)$ with respect to y . Thus

$$(2) \quad \frac{\partial f}{\partial y} \Big|_{(x_1, y_1)} = \frac{\partial z}{\partial y} \Big|_{(x_1, y_1)} = \lim_{\Delta y \rightarrow 0} \frac{f(x_1, y_1 + \Delta y) - f(x_1, y_1)}{\Delta y}.$$

Other notations are $f_y(x_1, y_1)$ and $z_y(x_1, y_1)$.

It is clear that the values of those two partial derivatives will in general depend on the point (x_1, y_1) , and so the partial derivatives $\partial z/\partial x$ and $\partial z/\partial y$ at a variable point (x, y) are functions of x and y . The function $\partial z/\partial x$ is obtained as in ordinary calculus by differentiating $z = f(x, y)$ with respect to x , treating y as a constant, and $\partial z/\partial y$ is obtained by differentiating z with respect to y , treating x as a constant.

Example 1. Let $z = f(x, y) = x^2y + x \sin y$. Then

$$\frac{\partial f}{\partial x} = 2xy + \sin y, \quad \frac{\partial f}{\partial y} = x^2 + x \cos y.$$

The partial derivatives $\partial z/\partial x$ and $\partial z/\partial y$ of a function $z = f(x, y)$ have a very simple geometric interpretation. The function $z = f(x, y)$ can be represented by a surface in space. The equation $y = y_1$ then represents a vertical plane intersecting the surface in a curve, and the partial derivative $\partial z/\partial x$ at a point (x_1, y_1) is the slope of the tangent (i.e., $\tan \alpha$ where α is the angle shown in Fig. 13) to the curve. Similarly, the partial derivative $\partial z/\partial y$ at (x_1, y_1) is the slope of the tangent to the curve $x = x_1$ on the surface $z = f(x, y)$ at (x_1, y_1) .

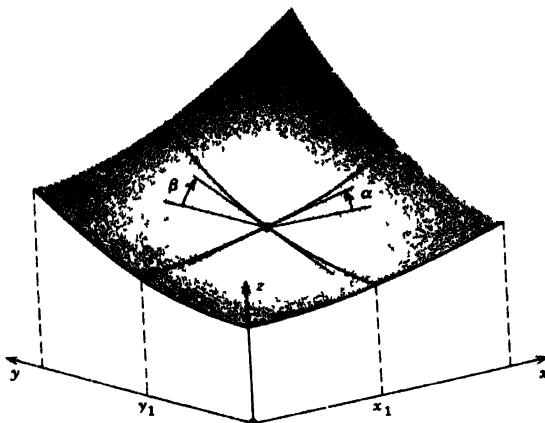


Fig. 13. Geometrical interpretation of first partial derivatives.

The partial derivatives $\partial z / \partial x$ and $\partial z / \partial y$ are called *first partial derivatives* or *partial derivatives of the first order*. By differentiating these derivatives once more we obtain the four *second partial derivatives* (or *partial derivatives of the second order*):

$$(3) \quad \frac{\partial^2 z}{\partial x^2} = f_{xx}(x, y), \quad \frac{\partial^2 z}{\partial x \partial y} = f_{xy}(x, y), \quad \frac{\partial^2 z}{\partial y \partial x} = f_{yx}(x, y), \quad \frac{\partial^2 z}{\partial y^2} = f_{yy}(x, y).$$

Thus $\partial^2 z / \partial x^2$ is the partial derivative of $\partial z / \partial x$ with respect to x , $\partial^2 z / \partial x \partial y$ is the partial derivative of $\partial z / \partial x$ with respect to y , etc.

It can be shown¹ that if all the derivatives concerned are continuous, then

$$(4) \quad \frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x},$$

i.e., the order of differentiation is then immaterial.

Example 2. For the function in Ex 1,

$$f_{xx} = 2y, \quad f_{xy} = 2x + \cos y = f_{yx}, \quad f_{yy} = -x \sin y.$$

By differentiating the second partial derivatives again with respect to x and y , respectively, we obtain the *third partial derivatives* or *partial derivatives of the third order* of f , etc.

If we consider a function $f(x, y, z)$ of three independent variables, then we have the three first partial derivatives $f_x(x, y, z)$, $f_y(x, y, z)$, and $f_z(x, y, z)$. Here f_x is obtained by differentiating f with respect to x , *treating both y and z as constants*. Thus, in analogy to (1), we now have

$$\frac{\partial f}{\partial x} \Big|_{(x_1, y_1, z_1)} = \lim_{\Delta x \rightarrow 0} \frac{f(x_1 + \Delta x, y_1, z_1) - f(x_1, y_1, z_1)}{\Delta x}$$

¹Cf. Ref. [A1] in Appendix 1.

etc. By differentiating f_x , f_y , f_z again in this fashion we obtain the second partial derivatives of f , etc.

Example 3. Let $f(x, y, z) = x^2 + y^2 + z^2 + xyx^2$. Then

$$\begin{array}{lll} f_x = 2x + ye^x & f_y = 2y + xe^x, & f_z = 2z + xye^x, \\ f_{xx} = 2, & f_{xy} = f_{yx} = e^x, & f_{xz} = f_{zx} = ye^x, \\ f_{yy} = 2, & f_{yz} = f_{zy} = xe^x, & f_{zz} = 2 + xye^x. \end{array}$$

Problems

Find the first partial derivatives of the following functions.

1. $f(x, y) = \sqrt{x^2 + y^2}$
2. $f(r, \theta) = r \cos \theta$
3. $f(x, y) = \arctan \frac{y}{x}$
4. $f(x, y) = e^{xy}$
5. $f(x, y) = \frac{1}{\sqrt{(x - a)^2 + (y - b)^2}}$ (a, b constant)
6. $f(r, h) = \frac{\pi}{3} r^2 h$
7. $f(R_1, R_2) = \frac{R_1 + R_2}{R_1 R_2}$
8. Find $f_x + g_y$, where $f(x, y) = x^2 - y^2$ and $g(x, y) = 2xy$.
9. Find $f_x + g_y$, where $f(x, y) = \ln(x^2 + y^2)$ and $g(x, y) = 2 \arctan \frac{y}{x}$.

Find $f_{xx} + f_{yy}$ where

$$10. f(x, y) = x^2 - y^2 \quad 11. f(x, y) = e^x \cos y \quad 12. f(x, y) = \sin x \cosh y$$

Sketch the surfaces corresponding to the following functions.

13. $z = x^2 + y^2$
14. $z = \ln(x^2 + y^2)$
15. $z = e^{xy}$
16. The curves $z = f(x, y) = \text{const}$ are called **level curves** of $f(x, y)$. Draw the level curves of the functions in Probs. 13–15.

Find the first partial derivatives of the following functions at the given points.

$$17. f(x, y) = \sqrt{1 - x^2 - y^2}, \text{ at } (0, 0) \quad 18. f(x, y) = (x^2 + y^2)^2, \text{ at } (1, 2)$$

Find the first and second partial derivatives of the following functions.

$$19. f(x, y, z) = \frac{1}{x^2 + y^2 + z^2} \quad 20. f(r, \theta, z) = \frac{x + y + z}{xyz}$$

0.3 Second and Third Order Determinants

Consider the system

$$(1) \quad \begin{aligned} a_1x + b_1y &= k_1 \\ a_2x + b_2y &= k_2 \end{aligned}$$

consisting of two linear equations in the unknowns x and y . To solve this system we may multiply the first equation by b_2 , the second by $-b_1$, and add, finding

$$(a_1b_2 - a_2b_1)x = k_1b_2 - k_2b_1.$$

Then we multiply the first equation of (1) by $-a_2$, the second by a_1 , and add again, finding

$$(a_1 b_2 - a_2 b_1)y = a_1 k_2 - a_2 k_1.$$

If $a_1 b_2 - a_2 b_1$ is not zero, we may divide and obtain the desired result

$$(2) \quad x = \frac{k_1 b_2 - k_2 b_1}{a_1 b_2 - a_2 b_1}, \quad y = \frac{a_1 k_2 - a_2 k_1}{a_1 b_2 - a_2 b_1}$$

The expression in the denominators is written in the form

$$\begin{array}{cc} a_1 & b_1 \\ a_2 & b_2 \end{array}$$

and is called a **determinant of the second order**; thus

$$(3) \quad \begin{array}{cc} a_1 & b_1 \\ a_2 & b_2 \end{array} = a_1 b_2 - a_2 b_1.$$

The four numbers a_1, b_1, a_2, b_2 are called the **elements** of the determinant. The elements in a horizontal line are said to form a **row** and the elements in a vertical line are said to form a **column** of the determinant.

We may now write the solution (2) of the system (1) in the form

$$(4) \quad x = \frac{D_1}{D}, \quad y = \frac{D_2}{D} \quad (D \neq 0)$$

where

$$D = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad D_1 = \begin{vmatrix} k_1 & b_1 \\ k_2 & b_2 \end{vmatrix}, \quad D_2 = \begin{vmatrix} a_1 & k_1 \\ a_2 & k_2 \end{vmatrix}$$

The formula (4) is called **Cramer's rule**.² Note that D_1 is obtained by replacing the first column of D by the column with elements k_1, k_2 , and D_2 is obtained by replacing the last column of D by that column.

Each equation of the system (1) represents a straight line in the xy -plane, and a pair of numbers (x, y) is a solution of (1) if, and only if, the point P with coordinates x, y lies on both lines. Hence there are three possible cases:

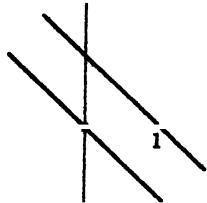
- (a) No solution if the lines are parallel.
- (b) Precisely one solution if they intersect.
- (c) Infinitely many solutions if they coincide.

² GABRIEL CRAMER (1704–1752), Italian mathematician, also known by his contributions to the theory of curves.

Example 1.

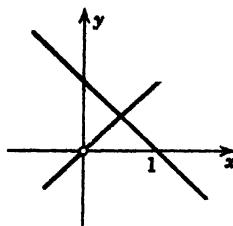
$$\begin{aligned}x + y &= 1 \\x + y &= 0\end{aligned}$$

Case (a)



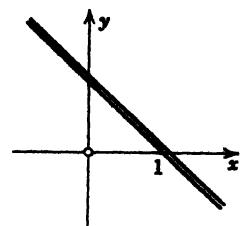
$$\begin{aligned}x + y &= 1 \\x - y &= 0\end{aligned}$$

Case (b)



$$\begin{aligned}x + y &= 1 \\2x + 2y &= 2\end{aligned}$$

Case (c)



If both k_1 and k_2 are zero, the system is said to be **homogeneous**; otherwise it is said to be **nonhomogeneous**.

If the system is homogeneous, Case (a) cannot occur, because then the lines represented by the equations pass through the origin, and the system has at least the *trivial solution* $x = 0, y = 0$.

The homogeneous system will have further solutions if, and only if, those two lines coincide, and then each point on the line is a solution. This happens if, and only if, $D = 0$, as the reader may show.

If the system is nonhomogeneous and $D \neq 0$, it has precisely one solution, which is obtained from (4).

A system of three linear equations in three unknowns x, y, z

$$(5) \quad \begin{aligned}a_1x + b_1y + c_1z &= k_1 \\a_2x + b_2y + c_2z &= k_2 \\a_3x + b_3y + c_3z &= k_3\end{aligned}$$

may be considered in a similar fashion. To obtain an equation involving only x the equations are multiplied, respectively, by

$$b_2c_3 - b_3c_2, \quad -(b_1c_3 - b_3c_1), \quad b_1c_2 - b_2c_1.$$

We see that these expressions may be written as second-order determinants:

$$M_1 = \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}, \quad -M_2 = -\begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix}, \quad M_3 = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}$$

Adding the resulting equations, we obtain

$$(6) \quad (a_1M_1 - a_2M_2 + a_3M_3)x = k_1M_1 - k_2M_2 + k_3M_3.$$

Two further equations containing only y and z , respectively, may be obtained in a similar manner.

To simplify our notation we now define a determinant of the third order by the equation

$$(7) \quad D = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}$$

We see that

$$D = a_1M_1 - a_2M_2 + a_3M_3,$$

the coefficient of x in (6), and if we write the second-order determinants in (7) at length, we obtain

$$D = a_1(b_2c_3 - b_3c_2) - a_2(b_1c_3 - b_3c_1) + a_3(b_1c_2 - b_2c_1)$$

or

$$(8) \quad D = a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1.$$

Obviously the determinant on the right-hand side of (7) which is multiplied by a_i , $i = 1, 2$, or 3 , is obtained from D by omitting the first column and the i th row of D .

We see that (6) may now be written

$$Dx = D_1$$

where

$$D_1 = \begin{vmatrix} k_1 & b_1 & c_1 \\ k_2 & b_2 & c_2 \\ k_3 & b_3 & c_3 \end{vmatrix}.$$

The aforementioned equation containing only y may be written

$$Dy = D_2$$

where

$$D_2 = \begin{vmatrix} a_1 & k_1 & c_1 \\ a_2 & k_2 & c_2 \\ a_3 & k_3 & c_3 \end{vmatrix},$$

and the equation containing only z may be written

$$Dz = D_3$$

where

$$D_3 = \begin{vmatrix} a_1 & b_1 & k_1 \\ a_2 & b_2 & k_2 \\ a_3 & b_3 & k_3 \end{vmatrix}.$$

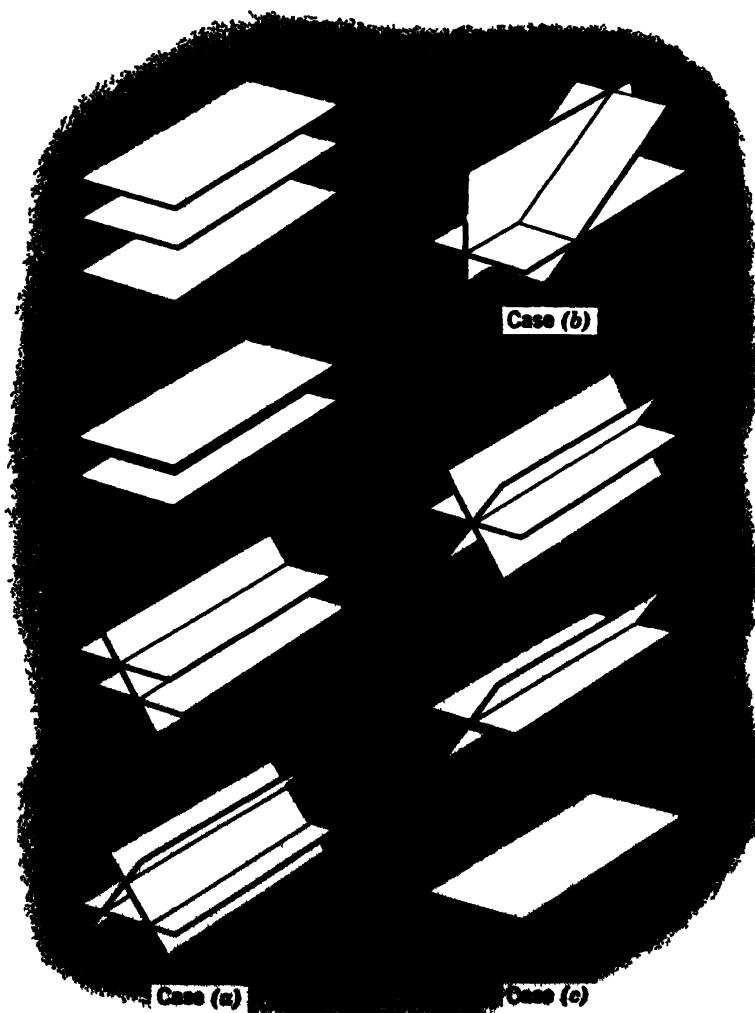
Note that the elements of D are arranged in the same order as they occur as coefficients in the equations of (5), and D_j , $j = 1, 2$, or 3 , is obtained from D by replacing the j th column by the column with elements k_1, k_2, k_3 , the expressions on the right sides of the equations of (5).

It follows that if $D \neq 0$, then system (5) has the unique solution

$$(9) \quad \frac{D_1}{D} \quad y = \frac{D_2}{D} \quad , \quad \frac{D_3}{D} \quad (\text{Cramer's rule}).$$

Each equation of system (5) represents a plane in space, and a triple of numbers (x, y, z) is a solution of (5) if, and only if, the point P with coordinates x, y, z is a common point of the three planes. As before we have three possible cases:

- (a) No solution if two (or all three) planes are parallel, if two of them coincide and the third is parallel, or if the intersections of each pair of them are three parallel lines.
- (b) Precisely one solution, given by (9), if the planes have just one point in common.
- (c) Infinitely many solutions if they have a line in common or coincide.



p. 14. Geometrical interpretation of three linear equations in three unknowns.

If (5) is homogeneous, i.e., $k_1 = k_2 = k_3 = 0$, it has at least the trivial solution $x = y = z = 0$, and nontrivial solutions exist if, and only if, $D = 0$.

If (5) is nonhomogeneous and $D \neq 0$, it has precisely one solution which is obtained from (9).

We shall now list the most important properties of our determinants; the corresponding proofs follow from (7) by direct calculation.⁸

(A) The value of a determinant is not altered if its rows are written as columns in the same order,

$$(10) \quad \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

(B) If any two rows (or two columns) of a determinant are interchanged, the value of the determinant is multiplied by -1 . Example:

$$(11) \quad \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = - \begin{vmatrix} c_1 & c_2 & c_3 \\ b_1 & b_2 & b_3 \\ a_1 & a_2 & a_3 \end{vmatrix} = \begin{vmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

The second-order determinant obtained from

$$D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

by deleting one row and one column is called the minor of the element which belongs to the deleted row and column. Example: The minors of a_2 and b_2 in D are

$$\begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} \quad \text{and} \quad \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix},$$

respectively, etc.

The cofactor of the element of D in the i th row and the k th column is defined as $(-1)^{i+k}$ times the minor of that element. Example: The cofactors of a_2 and b_2 are

$$- \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} \quad \text{and} \quad \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix},$$

respectively. Furthermore we see that we may write (7) in the form

$$D = a_1 C_1 + a_2 C_2 + a_3 C_3,$$

⁸ Determinants of arbitrary order n will be defined in Sec. 7.3, and we shall see that they have quite similar properties.

where C_i is the cofactor of a_i in D . From this and the properties (A) and (B) we obtain the following property.

(C) *The determinant D may be developed by any row or column, that is, it may be written as the sum of the three elements of any row (or column) each multiplied by its cofactor. For example, the development of D by its second row is*

$$D = -a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + b_2 \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}$$

From (C) we obtain

(D) *A factor of the elements of any row (or column) can be placed before the determinant. Example:*

$$\begin{vmatrix} 4 & 6 & 1 \\ 3 & -9 & 2 \\ -1 & 12 & 5 \end{vmatrix} = \begin{vmatrix} 4 & 2 \cdot 3 & 1 \\ 3 & -3 \cdot 3 & 2 \\ -1 & 4 \cdot 3 & 5 \end{vmatrix} = 3 \begin{vmatrix} 4 & 2 & 1 \\ 3 & -3 & 2 \\ -1 & 4 & 5 \end{vmatrix}$$

From the properties (B) and (D) we may draw the following conclusion.

(E) *If corresponding elements of two rows (or columns) of a determinant are proportional, the value of the determinant is zero.*

The following property is basic for simplifying determinants to be evaluated.

(F) *The value of a determinant remains unaltered if the elements of one row (or column) are altered by adding to them any constant multiple of the corresponding elements in any other row (or column). Example:*

$$\begin{aligned} \begin{vmatrix} -6 & 21 & -30 \\ 1 & -3 & 5 \\ 2 & 7 & -4 \end{vmatrix} &= \begin{vmatrix} -6 + 1 \cdot 7 & 21 - 3 \cdot 7 & -30 + 5 \cdot 7 \\ 1 & -3 & 5 \\ 2 & 7 & -4 \end{vmatrix} \\ &= \begin{vmatrix} 1 & 0 & 5 \\ 1 & -3 & 5 \\ 2 & 7 & -4 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 5 \\ 0 & -3 & 0 \\ 2 & 7 & -4 \end{vmatrix} = -3 \begin{vmatrix} 1 & 5 \\ 2 & -4 \end{vmatrix} = 42. \end{aligned}$$

(G) *If each element of a row (or column) of a determinant is expressed as a binomial, the determinant can be written as the sum of two determinants.*

Example:

$$\begin{vmatrix} a_1 + d_1 & b_1 & c_1 \\ a_2 + d_2 & b_2 & c_2 \\ a_3 + d_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} d_1 & b_1 \\ d_2 & b_2 \\ d_3 & b_3 \end{vmatrix}$$

By applying the product rule of differentiation we obtain the following property.

(H) If the elements of a determinant are differentiable functions of a variable, the derivative of the determinant may be written as a sum of three determinants,

$$\frac{d}{dx} \begin{vmatrix} f & g & h \\ p & q & r \\ u & v & w \end{vmatrix} = \begin{vmatrix} f' & g' & h' \\ p & q & r \\ u & v & w \end{vmatrix} + \begin{vmatrix} f & g & h \\ p' & q' & r' \\ u & v & w \end{vmatrix} + \begin{vmatrix} f & g & h \\ p & q & r \\ u' & v' & w' \end{vmatrix},$$

where primes denote derivatives with respect to x .

Determinants of higher order and more general systems of linear equations will be considered in Chap. 7.

Problems

Evaluate

1.
$$\begin{vmatrix} 3 & -1 \\ 2 & 4 \end{vmatrix}$$

2.
$$\begin{vmatrix} 0 & 3 \\ 5 & 7 \end{vmatrix}$$

3.
$$\begin{vmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{vmatrix}$$

4.
$$\begin{vmatrix} 4.6 & -4.1 \\ -2.0 & 0.2 \end{vmatrix}$$

5.
$$\begin{vmatrix} 4 & 12 \\ 40 & 20 \end{vmatrix}$$

6.
$$\begin{vmatrix} \sqrt{3} & -2 \\ 0.5 & \sqrt{27} \end{vmatrix}$$

7.
$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$$

8.
$$\begin{vmatrix} -4 & 18 & 7 \\ 0 & -1 & 4 \\ 0 & 0 & 6 \end{vmatrix}$$

9.
$$\begin{vmatrix} 9 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

10.
$$\begin{vmatrix} 6 & 13 & -2 \\ 5 & 37 & -5 \\ 1 & 13 & -2 \end{vmatrix}$$

11.
$$\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix}$$

12.
$$\begin{vmatrix} 17 & 4 & 6 \\ 8 & 0 & 5 \\ 4 & 0 & 3 \end{vmatrix}$$

13.
$$\begin{vmatrix} 1 & c & -b \\ -c & 1 & a \\ -b & -a & 1 \end{vmatrix}$$

14.
$$\begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix}$$

15.
$$\begin{vmatrix} a+b & b & 0 \\ b & b+c & c \\ 0 & c & c+d \end{vmatrix}$$

Solve the following systems of equations.

16.
$$\begin{aligned} 17x + 4y &= -24 \\ x - 3y &= 18 \end{aligned}$$

17.
$$\begin{aligned} 3x - y &= 1 \\ x + 3y &= 7 \end{aligned}$$

18.
$$\begin{aligned} 3x - 4y &= 14 \\ -x + 3y &= -8 \end{aligned}$$

19. Plot the straight lines represented by the equations in Prob. 17 and determine their point of intersection.

20. Show that the equation

$$\begin{vmatrix} y & x^2 & x \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{vmatrix} = 0$$

represents a parabola passing through the points $(0, 0)$, $(-1, 1)$, and $(1, 2)$.

0.4 Complex Numbers

The fact that there are equations such as

$$x^2 + 3 = 0, \quad x^2 - 10x + 40 = 0,$$

which are not satisfied by any real number, leads to the introduction of complex numbers.⁴ These are of the form $a + ib$ (or $a + bi$) where a and b are real numbers. The symbol i is called the **imaginary unit**; the number a is called the **real part** and b the **imaginary part** of $a + ib$. For example, the real part of $4 - 3i$ is 4, and the imaginary part is -3 . To each ordered pair of real numbers a, b there corresponds one complex number $a + ib$, and conversely.

Complex numbers can be represented as points in the plane. For this purpose we choose two perpendicular coordinate axes, the horizontal x -axis and the vertical y -axis and on both axes the same unit of length (Fig. 15).

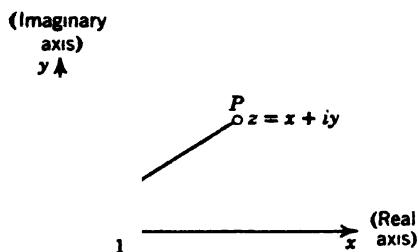


Fig. 15. The complex plane.

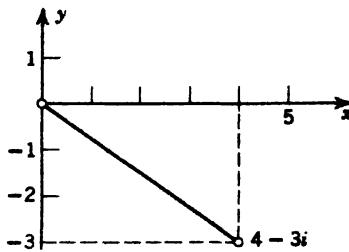


Fig. 16. The number $4 - 3i$ in the complex plane.

The xy -coordinate system thus obtained is called a **Cartesian coordinate system**.⁵ To the complex number $z = x + iy$ there corresponds the point P with Cartesian coordinates (x, y) . The xy -plane in which the complex numbers are represented geometrically in this fashion is called the **complex plane** or **Argand diagram**.⁶ The x -axis is called the **real axis**, and the y -axis the **imaginary axis**.

Two complex numbers

$$z_1 = x_1 + iy_1 \quad \text{and} \quad z_2 = x_2 + iy_2$$

⁴ First to use complex numbers for this purpose was the Italian mathematician GIROLAMO CARDANO (1501–1576), who found the formula for solving cubic equations. The term "complex number" was introduced by the great German mathematician CARL FRIEDRICH GAUSS (1777–1855) whose work was of basic importance in algebra, number theory, differential equations, differential geometry, non-Euclidean geometry, complex analysis, numerical analysis, and theoretical mechanics. He also paved the way for a general and systematic use of complex numbers.

⁵ Named after the French philosopher and mathematician RENATUS CARTESIUS (latinized for RENE DESCARTES (1596–1650)), who invented analytic geometry.

⁶ JEAN ROBERT ARGAND (1768–1822), French mathematician. His paper on the complex plane appeared in 1806, nine years after a similar memoir by the Norwegian mathematician CASPAR WESSEL (1745–1818).

are defined to be equal if, and only if, their real parts are equal and their imaginary parts are equal; that is,

$$z_1 = z_2 \quad \text{if, and only if,} \quad x_1 = x_2 \text{ and } y_1 = y_2.$$

If z_1 and z_2 are different (not equal), then we may write $z_1 \neq z_2$.

Inequalities between complex numbers, such as $z_1 < z_2$ or $z_1 \geq z_2$, have no meaning. (Inequalities may hold between the *absolute values* of complex numbers; see Sec. 0.5.)

Addition. The sum $z_1 + z_2$ of z_1 and z_2 is defined as the complex number obtained by adding the real parts and the imaginary parts of z_1 and z_2 , that is,

$$(1) \quad z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2).$$

We see that addition of complex numbers is in accordance with the “*parallelogram law*” by which forces are added in mechanics (Fig. 17).

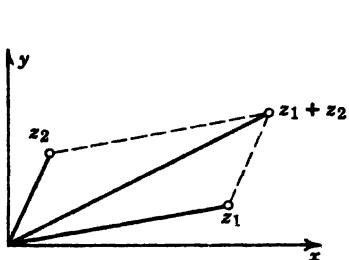


Fig. 17. Addition of complex numbers.

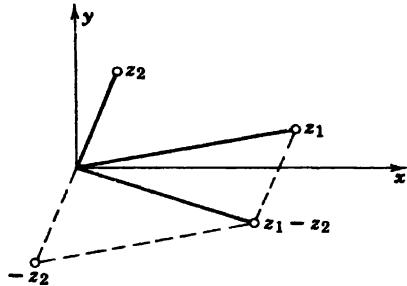


Fig. 18. Subtraction of complex numbers.

Subtraction. This operation is defined as the inverse operation of addition; that is, the difference $z_1 - z_2$ is the complex number z for which $z_1 = z + z_2$. Obviously,

$$(2) \quad z_1 - z_2 = (x_1 - x_2) + i(y_1 - y_2).$$

Multiplication. The product $z_1 z_2$ is defined as the complex number

$$(3) \quad z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1),$$

which is obtained formally by applying the ordinary rules of arithmetic for real numbers, treating the symbol i as a number, and replacing $i^2 = ii$ by -1 .

Division. This operation is defined as the inverse operation of multiplication; that is, the quotient $z = z_1/z_2$ is the complex number $z = x + iy$ for which

$$z_1 = zz_2 = (x + iy)(x_2 + iy_2).$$

By using (3) we see that this can be written

$$x_1 + iy_1 = (x_2 x - y_2 y) + i(y_2 x + x_2 y).$$

By definition of equality the real parts and the imaginary parts on both sides must be equal:

$$x_1 = x_2 x - y_2 y$$

$$y_1 = y_2 x + x_2 y.$$

This is a system of two linear equations in the unknowns x and y . The determinant of the coefficients has the value $x_2^2 + y_2^2$. Assuming that $z_2 = x_2 + iy_2 \neq 0$ that determinant is not zero, and Cramer's rule (Sec. 0.3) yields the unique solution

$$x = \frac{x_1 x_2 + y_1 y_2}{x_2^2 + y_2^2}, \quad y = \frac{x_2 y_1 - x_1 y_2}{x_2^2 + y_2^2} \quad (z_2 \neq 0).$$

From this we see that the number $z = x + iy$ can be obtained formally by multiplying both the numerator and the denominator of the quotient z_1/z_2 by $x_2 - iy_2$; thus

$$(4) \quad z = \frac{z_1}{z_2} = \frac{(x_1 + iy_1)(x_2 - iy_2)}{(x_2 + iy_2)(x_2 - iy_2)} = \frac{x_1 x_2 + y_1 y_2}{x_2^2 + y_2^2} + i \frac{x_2 y_1 - x_1 y_2}{x_2^2 + y_2^2}.$$

Example 1. Let $z_1 = 2 - 3i$ and $z_2 = -5 + i$. Then

$$z_1 + z_2 = -3 - 2i, \quad z_1 - z_2 = 7 - 4i, \quad z_1 z_2 = -7 + 17i,$$

$$\frac{z_1}{z_2} = \frac{(2 - 3i)(-5 - i)}{(-5 + i)(-5 - i)} = \frac{-10 - 2i + 15i - 3}{25 + 1} = -\frac{1}{2} + \frac{i}{2}.$$

For any complex numbers z_1, z_2, z_3 we have

$$\begin{aligned} z_1 + z_2 &= z_2 + z_1 && \text{(Commutative laws)} \\ z_1 z_2 &= z_2 z_1 \end{aligned}$$

$$(5) \quad \begin{aligned} (z_1 + z_2) + z_3 &= z_1 + (z_2 + z_3) && \text{(Associative laws)} \\ (z_1 z_2) z_3 &= z_1 (z_2 z_3) \end{aligned}$$

$$z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3 \quad \text{(Distributive law).}$$

These laws follow immediately from the corresponding laws for real numbers and the preceding definitions of the algebraic operations for complex numbers.

Let $z = x + iy$ be any complex number; then $x - iy$ is called the conjugate of z and is denoted by \bar{z} (Fig. 19). Thus

$$z = x + iy, \quad \bar{z} = x - iy.$$

For example, the conjugate of $4 + 7i$ is $4 - 7i$. Now

$$z + \bar{z} = 2x, \quad z - \bar{z} = 2iy$$

and therefore

$$(6) \quad \operatorname{Re} z = x = \frac{1}{2}(z + \bar{z}), \quad \operatorname{Im} z = y = \frac{1}{2i}(z - \bar{z})$$

where the symbol Re denotes the real part and Im denotes the imaginary part. Obviously,

$$\overline{(z_1 + z_2)} = \bar{z}_1 + \bar{z}_2, \quad (z_1 - z_2) = \bar{z}_1$$

$$\overline{(z_1 z_2)} = \bar{z}_1 \bar{z}_2, \quad \overline{\left(\frac{z_1}{z_2}\right)} = \frac{\bar{z}_1}{\bar{z}_2}.$$

In the fundamental operations of arithmetic the complex number $x + i0$ behaves like the real number x . Indeed, equality, sum, and product as defined for two such numbers $x_1 + i0$ and $x_2 + i0$ carry over, respectively, into equality, sum, and product as defined for two real numbers x_1 and x_2 . For this reason we may say that a complex number whose imaginary part is zero is a real number, and write simply x instead of $x + i0$. The complex number $0 + iy$ is written iy and is called a **pure imaginary number**. Defining i to mean $0 + i1$ we have

$$(8) \quad i^2 = (0 + i1)(0 + i1) = -1$$

and so i^2 may be replaced by -1 whenever it occurs. Consequently,

$$(9) \quad \begin{aligned} i^2 &= -1, & i^3 &= -i, & i^4 &= 1, & i^5 &= i, \dots \\ \frac{1}{i} &= \frac{i}{i^2} = -i, & \frac{1}{i^2} &= -1, \dots \end{aligned}$$

Problems

Let $z_1 = 3 - 2i$ and $z_2 = 2 + 4i$. Reduce each of the following expressions to the form $a + ib$.

1. $z_1 z_2, z_1/z_2, z_1^2, z_1^2/z_2, (1 + i)z_1$
2. $z_1 + z_2, z_1 - z_2, z_2 - z_1, z_2^3, (z_1/z_2)^2, z_2/(1 - 2i)$

Reduce each of the following expressions to the form $a + ib$.

3. $(1 + i)^2, \left(\frac{1 + i}{1 - i}\right)^2$
4. $\left(\frac{1 + i}{1 - i}\right)^2 - \left(\frac{1 - i}{1 + i}\right)^2$

Solve the following equations.

5. $z^2 + 9 = 0$
6. $z^2 - 2z + 2 = 0$
7. $z^2 + 2z + 5 = 0$
8. $z^2 + z + 9 = 0$

9. Prove the first two formulas of (7).

10. Prove the last two formulas of (7).

11. If $z = \bar{z}$, show that z is real. If $z = -\bar{z}$, show that z is pure imaginary or zero.

12. If $z^2 = \bar{z}^2$, show that z is either real or pure imaginary.

13. Prove that any number is equal to the conjugate of its conjugate.

14. Plot the numbers $2 + 4i, 2 - 4i, -2 + 4i, -2i$, and $1 - i$ as points in the complex plane.

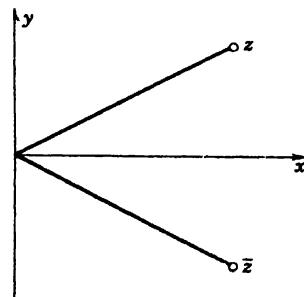


Fig. 19. Conjugate complex numbers.

15. If $z_1 z_2 = 0$, show that at least $z_1 = 0$ or $z_2 = 0$.

Let $z = x + iy$. Find

16. $\operatorname{Re}(1/z)$, $\operatorname{Im}(1/z)$

17. $\operatorname{Im}(z^3)$

18. $\operatorname{Re}(1/z^2)$, $\operatorname{Re}(z^2 + z)$

19. $\operatorname{Re}(-iz^2)$, $\operatorname{Im}(4iz^2 - 6z + 8i)$

20. $\operatorname{Re}[1/(z - i)]$

0.5 Polar Form of Complex Numbers

If we introduce polar coordinates r, θ in the complex plane by setting

$$(1) \quad x = r \cos \theta, \quad y = r \sin \theta,$$

then the complex number $z = x + iy$ may be written

$$(2) \quad z = r \cos \theta + ir \sin \theta = r(\cos \theta + i \sin \theta).$$

This is known as the **polar form** or *trigonometric form* of a complex number. r is called the **absolute value** or *modulus* of z and is denoted by $|z|$. Thus

$$(3) \quad |z| = r = \sqrt{x^2 + y^2} = \sqrt{z\bar{z}} \quad (\geq 0).$$

Obviously, r is the distance of the point P corresponding to z from the origin O in the complex plane (Fig. 20). The directed angle measured from the positive x -axis to OP is called the **argument** of z and is denoted by $\arg z$; *angles will be measured positive in the counterclockwise direction and in terms of radians*. We have, then,

$$(4) \quad \arg z = \theta = \operatorname{arc} \sin \frac{y}{r} = \operatorname{arc} \cos \frac{x}{r} = \operatorname{arc} \tan \frac{y}{x},$$

as in trigonometry. Note that for given z , the argument θ is determined only up to multiples of 2π .

Example 1. Let $z = 1 + i$. Then $|z| = \sqrt{2}$ and $\arg z = \frac{\pi}{4} + 2n\pi$ where $n = 0, 1, 2, \dots$

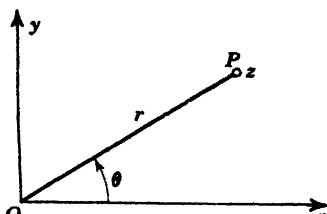


Fig. 20. Trigonometric form of complex numbers.

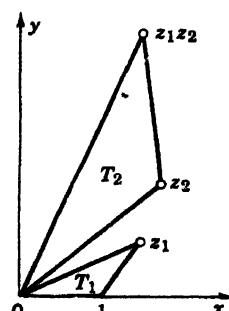


Fig. 21. Multiplication of complex numbers.

The polar form of complex numbers is particularly useful in analyzing multiplication and division. Let

$$z_1 = r_1(\cos \theta_1 + i \sin \theta_1) \quad \text{and} \quad z_2 = r_2(\cos \theta_2 + i \sin \theta_2).$$

Then the product is

$$z_1 z_2 = r_1 r_2[(\cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2) + i(\sin \theta_1 \cos \theta_2 + \cos \theta_1 \sin \theta_2)].$$

If we apply the familiar addition theorems of the sine and the cosine, this assumes the simple form

$$(5) \quad z_1 z_2 = r_1 r_2[\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)].$$

We thus obtain the important rules

$$(6) \quad |z_1 z_2| = |z_1| |z_2|$$

and

$$(7) \quad \arg(z_1 z_2) = \arg z_1 + \arg z_2 \quad (\text{up to multiples of } 2\pi).$$

Similarly, from the definition of division it follows that

$$(8) \quad \left| \frac{z_1}{z_2} \right| = \frac{|z_1|}{|z_2|}$$

and

$$(9) \quad \arg \frac{z_1}{z_2} = \arg z_1 - \arg z_2 \quad (\text{up to multiples of } 2\pi).$$

Table 2. arc tan :

$\frac{y}{x}$	arc tan $\frac{y}{x}$						
0	0.00000	1.0	0.78540	2.0	1.10715	4.0	1.32582
0.1	0.09967	1.1	0.83298	2.2	1.14417	4.5	1.35213
0.2	0.19740	1.2	0.87606	2.4	1.17601	5.0	1.37340
0.3	0.29146	1.3	0.91510	2.6	1.20362	5.5	1.39094
0.4	0.38051	1.4	0.95055	2.8	1.22777	6.0	1.40565
0.5	0.46365	1.5	0.98279	3.0	1.24905	7.0	1.42890
0.6	0.54042	1.6	1.01220	3.2	1.26791	8.0	1.44644
0.7	0.61073	1.7	1.03907	3.4	1.28474	9.0	1.46014
0.8	0.67474	1.8	1.06370	3.6	1.29985	10.0	1.47113
0.9	0.73282	1.9	1.08632	3.8	1.31347	11.0	1.48014

Example 2. From (6) and (7) we obtain

$$z^n = r^n(\cos \theta + i \sin \theta)^n = r^n(\cos n\theta + i \sin n\theta)$$

and from this so-called formula of De Moivre⁷

$$(10) \quad (\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta.$$

From (6) and (7) it follows that the triangles T_1 and T_2 in Fig. 21 are similar. This can be used as the basis for a graphical construction of the product of complex numbers.

⁷ABRAHAM DE MOIVRE (1667–1754), French mathematician, who introduced imaginary quantities in trigonometry and contributed to the theory of mathematical probability.

Problems

Represent the following complex numbers in polar form.

1. $2 - 2i, i, 3 + 4i$
2. $5 + 5i, -5 + 5i, -5 - 5i$
3. Show that $\arg \bar{z} = -\arg z$ (up to multiples of 2π).
4. Show that $\arg(1/z) = -\arg z$ (up to multiples of 2π).
5. Prove that $|a - z| = |a - \bar{z}|$ where a is any real number. Interpret geometrically.
6. Prove that $|z_1 - z_2|$ is the distance between the points z_1 and z_2 in the complex plane.

Find the absolute values of the following numbers.

7. $1 + i\sqrt{3}, -9i, 2 + i\sqrt{5}$
8. $2 - i\sqrt{5}, 2 + 3i, (4 + i)^3$
9. $\frac{(4 - 3i)(\frac{1}{2} + i)^4}{\left(\frac{1 - 3i}{4}\right)^2(-3 + 4i)}$
10. $\left(\frac{1 + i}{1 - i}\right)^8, (3 + 4i)^3(-1 - i)^6$

What loci are represented by the following equations and inequalities? (Plot a graph in each case.)

11. $|z| = 1$
12. $|z - 1| = 1$
13. $\operatorname{Re}(z^2) = -1$
14. $\operatorname{Im}(2z) = -1$
15. $0 \leq \arg z \leq \pi/2$
16. $0 \leq \arg(1/z) \leq \pi/2$
17. Using (10), prove the identities
 $\cos 3\theta = \cos^3 \theta - 3 \cos \theta \sin^2 \theta, \quad \sin 3\theta = 3 \cos^2 \theta \sin \theta - \sin^3 \theta.$
18. Find similar identities as in Prob. 17 for $\cos 2\theta$ and $\cos 4\theta$.
19. Find $(3 - i)(2 + i)$ by means of a graphical construction in the complex plane.
20. For given z find $1/z$ by a graphical construction.

0.6 Some General Remarks About Numerical Computations

The present section is devoted to a few simple but important general remarks about numerical computations.

It is clear that engineering mathematics ultimately comes down to numerical results, and the engineering student should, therefore, supplement his mathematical equipment with a definite knowledge of some fundamental numerical methods.

Various examples and problems in the text will help the student to learn how to arrange computations in a suitable form. However, numerical computation requires practical experience. It cannot be learned solely from books; practical training is needed, just as in swimming, driving a car, or playing the piano. Consequently, the student should not only work out book examples and problems, but also set up and calculate examples of his own. Active work is more important in numerical analysis than in many other branches of mathematics. Famous mathematicians like Gauss spent a considerable part of their time with numerical computations, and the engineering student will do well in developing a similar attitude.

In many cases formally complete answers to a problem obtained by theoretical considerations may be almost useless for numerical purposes. For example, in the case of systems of linear equations we obtain the solution in the form of quotients of determinants (Secs. 0.3 and 7.5), but if the number of equations and unknowns is large, the direct evaluation of the solution in this form is certainly not the *practical* answer to the problem of finding *numerical values* of the unknowns. In other cases, the theory may yield the solution of a problem in the form of an integral which cannot be evaluated by the usual methods of calculus. Then we need an approximation method for obtaining numerical values of that solution. Moreover, there are many practical problems which cannot be solved by exact methods at all, and in such cases, which arise quite frequently, we have to look for an appropriate approximation method which will yield numerical values of the solution of the problem.

Because we work with a finite number of digits and carry out a finite number of steps in computations, the methods of numerical analysis are *finite processes*, and a numerical result is an approximate value of the (unknown) exact result, except for the rare cases where the exact answer is a sufficiently simple rational number. The difference

$$E = a^* - a$$

between the exact value a and the approximate value a^* is called the *absolute error* of a^* , and the quotient

$$E_r = \frac{a^* - a}{a} \quad (a \neq 0)$$

is called the *relative error* of a^* .

An error may arise by the process of *rounding off*, that is, by retaining only the more significant decimal digits of a number and rejecting the less significant beyond a certain point. This error is called a **rounding error**. Rounding errors can often be eliminated by carrying one, two, or even more extra figures, known as *guarding figures*, in the intermediate steps of calculation.

Another error may result from the fact that one or several of the formulas used in numerical work are obtained by replacing an infinite series by one of its partial sums. Such an error is called a **truncation error**. It occurs, for instance, in connection with numerical integration and differentiation.

Moreover, a computation may contain **mistakes** because of the fallibility of the human computer or the mechanical and electrical equipment used in the computation. While errors cannot be avoided, mistakes are avoidable in principle. **Checking** of numerical results is quite important and any numerical method should include checking procedures to confirm that the results contain no mistakes. This certainly applies to final results, but in a longer computation it is also advisable to check intermediate results.

Of course, *any computation should be arranged in a systematic tabular form and should contain all the intermediate results directly on the working sheet.* This will help to avoid mistakes and to locate and correct mistakes that have been made. *Numerical work should not be done on scraps of paper.* The numbers should be written in a neat and legible manner on paper that is full notebook page size.

Our last remark is concerned with approximation methods. If a problem cannot be solved by an exact method and we therefore use an approximation method, the immediate basic question is how much at most can the approximate result deviate from the (unknown) exact result. This requires the *estimation of error* involved in approximation methods. In many cases formulas for estimating the maximum possible error are known, and they should be used whenever such an approximation formula is applied. This is the only way to obtain a clear picture of the quality of an approximation and to eliminate a source of stupid mistakes which otherwise may, and actually do, occur in the mathematical part of engineering work. The determination of new and more effective methods of estimating errors is one of the important and interesting problems in modern applied mathematics. The most recent developments in this direction are influenced considerably by the increasing use of electronic computers. Older methods were brought into forms suitable for electronic computation, and new methods have been added; this development is still in progress.

0.7 *Solution of Equations*

A frequent task in engineering mathematics is the determination of the roots of an equation of the form

$$(1) \quad f(x) = 0 \quad (x \text{ real}).$$

If $f(x)$ is a simple function, then there may be a formula which yields the exact values of the roots. For example, if

$$f(x) = ax^2 + bx + c \quad (a \neq 0)$$

we have to solve a *quadratic equation*, the solutions being

$$(2) \quad x = \frac{1}{2a} (-b \pm \sqrt{b^2 - 4ac}).$$

However, if $f(x)$ is more complicated, there may not be a formula for the exact values of the roots and in such a case we may use an approximation method which yields approximate values of the roots. Let us explain an important method of this type.

Suppose we want to determine the real roots of the equation $f(x) = 0$ where f is a differentiable function. Then we may plot a graph of $f(x)$ in the usual way. Assuming that the equation has a real root, we may obtain

from the graph a first rough approximation of this root; let x_1 be this approximation. We may then determine a second approximation x_2 , namely, the point of intersection of the tangent to the curve of $f(x)$ at x_1 and the x -axis (Fig. 22). Since

$$\tan \alpha = f'(x_1) = \frac{f(x_1)}{x_1 - x_2}$$

by solving for x_2 we obtain

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

In the next step we compute

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}$$

etc., and in general

$$(3) \quad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (n = 1, 2, \dots)$$

This method of successive approximations is called **Newton's⁸ method**.

Example 1. Given the equation

$$f(x) = x^3 + x - 1 = 0$$

we see that it has a root near $x_1 = 1$ (Fig. 23). From (3) we obtain

$$x_{n+1} = x_n - \frac{x_n^3 + x_n - 1}{3x_n^2 + 1} = \frac{2x_n^3 + 1}{3x_n^2 + 1} = \frac{N_n}{D_n}.$$

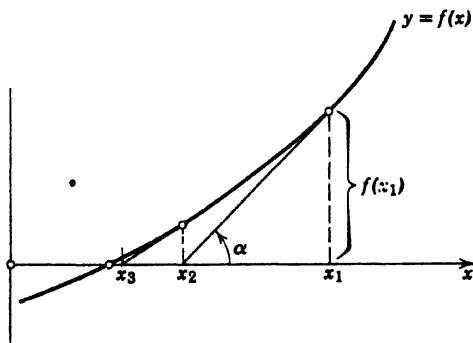


Fig. 22. Newton's method.

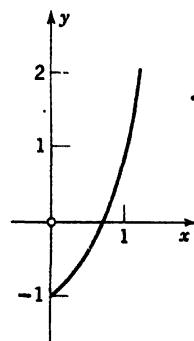


Fig. 23. Example 1.

⁸ Sir ISAAC NEWTON (1642–1727), English physicist and mathematician. Newton and the German mathematician and philosopher, GOTTFRIED WILHELM LEIBNIZ (1646–1716), invented (independently) the differential and integral calculus. Newton discovered many basic physical laws and introduced the method of investigating physical problems by means of calculus. His work is of greatest importance to both mathematics and physics.

The computation may be arranged in tabular form:

n	x_n	x_n^2	x_n^3	N_n	D_n	x_{n+1}
1	1	1	1	3	4	0.75
2	0.75	0.562 5	0.421 875	1.843 750	2.687 500	0.686 047
3	0.686 047	0.470 660	0.322 895	1.645 790	2.411 980	0.682 340
4	0.682 340	0.465 588	0.317 689	1.635 378	2.396 764	0.682 328

The result seems to indicate that the desired root to four decimal places is $x = 0.6823$. (We mention that even all six decimals of x_5 are the correct first decimals of the root.)

Sometimes work may be saved by starting with a small number of digits and increasing it gradually. Our present computation may then become:

n	x_n	x_n^2	x_n^3	N_n	D_n	x_{n+1}
1	1	1	1	3	4	0.75
2	0.75	0.562 5	0.421 9	1.843 8	2.687 5	0.686
3	0.686	0.470 60	0.322 83	1.645 66	2.411 80	0.682 3
4	0.682 3	0.465 533	0.317 633	1.635 266	2.396 599	0.682 33

Example 2. Find the positive root of the equation

$$2 \sin x = x.$$

Here $f(x) = x - 2 \sin x = 0$, and from (3),

$$x_{n+1} = x_n - \frac{x_n - 2 \sin x_n}{1 - 2 \cos x_n} = \frac{2(\sin x_n - x_n \cos x_n)}{1 - 2 \cos x_n} = \frac{N_n}{D_n}.$$

From the graph we conclude that the root is near $x_1 = 2$. Using tables of the sine and cosine functions, the computation yields

n	x_n	N_n	D_n	x_{n+1}
1	2	3.48	1.83	1.90
2	1.90	3.121 1	1.646 6	1.895 5
3	1.895 5	3.104 925 6	1.638 055 9	1.895 494

The root to four decimal places is $x = 1.8955$.

Newton's method approximates the curve of $f(x)$ by tangents. Complications will arise if the derivative $f'(x)$ is zero at or near a root of $f(x) = 0$, because then the denominator in (3) is small in absolute value. (For details see Ref. [A11] in Appendix 1.)

We shall now consider the so-called **regula falsi** (*method of false position*) which approximates that curve by a chord, as shown in Fig. 25. The chord intersects the x -axis at the point

$$(4) \quad x = \frac{af(b) - bf(a)}{f(b) - f(a)}$$

and this is an approximation for the root X_0 of the equation $f(x) = 0$.

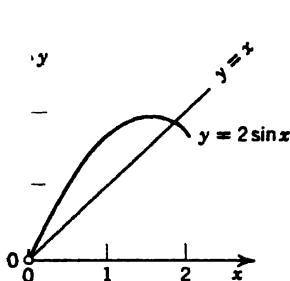


Fig. 24. Example 2.

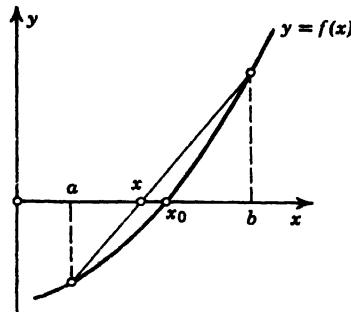


Fig. 25. Regula falsi.

Example 3. Find an approximate value of the root of the equation

$$f(x) = x^3 + x - 1 = 0$$

near $x = 1$ (cf. Ex. 1). We find that

$$f(0.5) = -0.375 \quad \text{and} \quad f(1) = 1.$$

Choosing $a = 0.5$ and $b = 1$, formula (4) yields

$$x_1 = \frac{0.5 \cdot 1 - 1 \cdot (-0.375)}{1 - (-0.375)} = 0.64.$$

Since $f(0.64) = -0.0979 < 0$, in the next step we may choose $a = 0.64$ and $b = 1$, and obtain from (4) the more accurate approximation $x_2 = 0.672$, and so on.

A third important method for determining the roots of an equation $f(x) = 0$ is the so-called **iteration method**.

This method is applicable if the given equation $f(x) = 0$ can be written in the form

$$x = g(x).$$

We choose a rough approximation x_1 of the root X_0 of the given equation and compute a second approximation

$$x_2 = g(x_1),$$

a third approximation

$$x_3 = g(x_2),$$

etc., and in general

$$(5) \qquad \qquad \qquad x_{n+1} = g(x_n) \qquad \qquad (n = 1, 2, \dots).$$

It can be shown that if

$$|g'(X_0)| < 1$$

and x_1 is sufficiently accurate (sufficiently close to X_0), then $x_n \rightarrow X_0$ as $n \rightarrow \infty$, that is, (5) will yield better and better approximations as n increases.

Example 4. The equation

(cf. Ex. 1) may be written

$$f(x) = x^3 + x - 1 = 0$$

$$(6) \quad x = \frac{1}{1 + x^2}$$

Hence, in this case, (5) takes the form

$$\frac{1}{1 + x_n^2}$$

Starting from $x_1 = 1$, as in Ex. 1, we thus obtain

$$x_2 = \frac{1}{1 + 1^2} = \frac{1}{2} = 0.5, \quad x_3 = \frac{1}{1 + (1/2)^2} = \frac{4}{5} = 0.8, \quad x_4 = 0.61, \text{ etc.}$$

Note that the right-hand side of (6) is

$$g(x) = \frac{1}{1 + x^2}, \quad \text{and} \quad g'(x) = \frac{-2x}{(1 + x^2)^2}.$$

Now $X_0 \approx 0.7$, and $|g'(0.7)| = 1.4/1.49^2 < 1$, which shows that for x_1 sufficiently close to X_0 we have $x_n \rightarrow X_0$ as $n \rightarrow \infty$. Our numerical results seem to confirm this fact.

It is clear that a given equation $f(x) = 0$ may be written in the form $x = g(x)$ in many ways. In the present example we may write

$$x = 1 - x^3$$

but then

$$g(x) = 1 - x^3, \quad |g'(x)| = 3x^2 > 1 \text{ when } x = X_0 \approx 0.7,$$

and the iteration will not work. The reader may try to start from $x_1 = 1, x_1 = 0.7, x_1 = 2$, and see what happens.

Problems

Find a root of the following equations; carry out three steps of Newton's method, starting from the given x_1 .

1. $x^3 - 1.2x^2 + 2x - 2.4 = 0, x_1 = 1$
2. $x^3 - 1.2x^2 + 2x - 2.4 = 0, x_1 = 2$
3. $x^3 - 3.9x^2 + 0.9x + 5.8 = 0, x_1 = 3$
4. $x^3 - 3.9x^2 + 0.9x + 5.8 = 0, x_1 = 1$
5. $x^3 - 3.9x^2 + 4.79x - 1.881 = 0, x_1 = 1$

6. The roots of the equation in Prob. 5 are 0.9, 1.1, and 1.9. A close to 0.9 and 1.1, Newton's method does not yield one of these roots. Why? Choose another x_1 such that the method yields approximations for the root 1.1.

Find all real roots of the following equations by Newton's method.

7. $x^4 - 0.1x^3 - 0.82x^2 - 0.1x - 1.82 = 0$
8. $x^3 - 4x^2 + 2x - 8 = 0$
9. $x^3 - 5x^2 + 6.64x - 1.92 = 0$
10. $\cos x = x$
11. $x + \ln x = 2$
12. $2x + \ln x = 1$

Find the real roots of the following equations by means of the regula falsi.

13. $x^4 = 2$
14. $x^4 = 2x$
15. $3 \sin x = 2x$

16. In Prob. 13 the approximate values of the positive root will always be somewhat smaller than the exact values of the root. Why?
17. Find the root of $x^5 = x + 0.2$ near $x = 0$ by the iteration method, starting from $x_1 = 0$.
18. The equation in Prob. 17 has a root near $x = 1$. Find this root by the iteration method, starting from $x_1 = 1$. Hint: Write the equation in the form $x = \sqrt[5]{x + 0.2}$.
19. What happens in Prob. 18 if you write the equation in the form $x = x^5 - 0.2$ and start from $x_1 = 1$?
20. Using the iteration method, show that the smallest positive root of the equation $x = \tan x$ is 4.49, approximately. Hint: Conclude from the graphs of x and $\tan x$ that a root lies close to $x_1 = 3\pi/2$; write the equation in the form $x = \arctan x$. (Why?)

0.8 Approximate Integration

If $f(x)$ is a given function and we can find a function $F(x)$ whose derivative is $f(x)$, then we may evaluate definite integrals of $f(x)$ by using the familiar formula

$$(1) \quad \int_a^b f(x) dx = F(b) - F(a) \quad [F'(x) = f(x)].$$

Tables of integrals⁹ may be helpful for this purpose. However, in engineering applications there frequently occur integrals which cannot be evaluated by familiar methods known from calculus. These may be integrals of two types, namely, integrals such as

$$\int e^{-x^2} dx, \quad \int \frac{\sin x}{x} dx, \quad \int \cos(x^2) dx,$$

which cannot be represented in terms of finitely many elementary functions, or integrals whose integrands are empirical functions, given by a table of numerical values which are obtained from a physical experiment or in some other way. In many cases, integrals of the first type may be evaluated by one of the advanced methods (complex integration, use of power series, or asymptotic expansions), and in the case of an integral of the second type we may try to approximate the empirical function by a polynomial or some other elementary function.

Nevertheless, in various situations it will be preferable to apply one of the standard methods of numerical integration which we shall consider in the present section. These methods use the fact that the definite integral in (1) equals the area A of the shaded region R in Fig. 26 on p. 32.

To determine A we may cut R from cardboard, weight the piece, and divide the result by the weight of a square of this cardboard whose side is 1. Another simple method is to draw R on a graph paper and count squares.

⁹ Cf. for example, Ref. [A3] in Appendix 1.

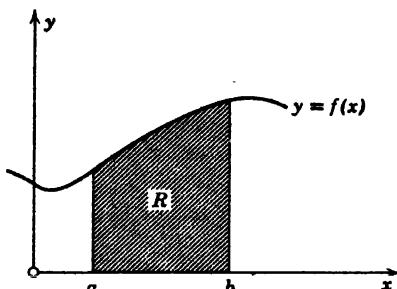


Fig. 26. Geometrical interpretation of a definite integral.

More accurate results are obtained by using a **planimeter**. So let us briefly discuss this instrumental method. Planimeters are inexpensive precision instruments which measure the area of any region R bounded by a closed curve C . There are several types of planimeters.¹⁰ Figure 27 shows a so-called *polar planimeter*. This instrument has a rod, the tracing arm (I), on one end of which is mounted the tracing pin (II) which follows the curve C . The other end of the tracing arm is moved on a circular path by means of a second rod, the pole arm (III), which can be turned about the end point 0 (the pole) and is joined to the other end of the tracing arm by means of a hinge. An integrating wheel (IV), mounted on the tracing arm, measures the area of the region R when the pin traces once around the entire boundary curve C , starting from any point of C and returning to that point; the area is obtained as the difference of the initial and final reading on the scale of the integrating wheel. The accuracy can be increased by repeating this procedure m times, adding the m results, and dividing their sum by m .

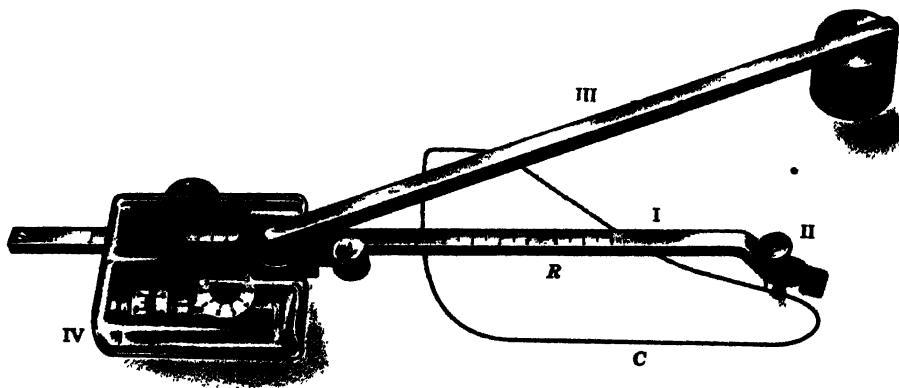


Fig. 27. Polar planimeter, made by A. Ott, Kempten, Germany, manufacturers of high precision instruments.

¹⁰ These types and the corresponding mathematical theories are considered in Ref. [A15] in Appendix 1.

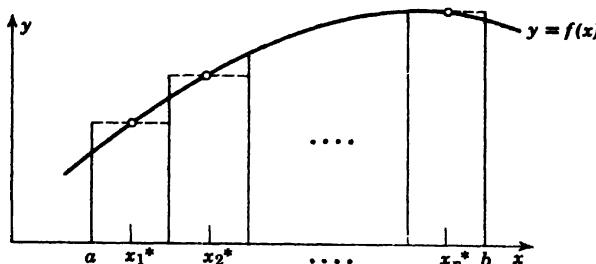


Fig. 28. Rectangular rule.

We shall now consider the simplest formulas of numerical integration, assuming that $b > a$ in (1).

We subdivide the interval of integration, whose length is $b - a$, into n equal parts of length

$$\Delta x = \frac{b - a}{n}$$

Let x_1^*, \dots, x_n^* be the midpoints of these n intervals. Then the n rectangles in Fig. 28 have the areas $f(x_1^*) \Delta x, \dots, f(x_n^*) \Delta x$. Therefore,

$$(2) \quad \int_a^b f(x) dx \approx \Delta x [f(x_1^*) + f(x_2^*) + \dots + f(x_n^*)].$$

This simple formula is called the **rectangular rule**.

Example 1. Use (2) to find an approximate value of $\ln 2$ ($= 0.693\ 15$). We have

$$\ln 2 = \int_1^2 \frac{dx}{x}$$

We choose $n = 5$ and $n = 10$ and arrange the work in tabular form.

j	x_j^*	$1/x_j^*$	j	x_j^*	$1/x_j^*$	j	x_j^*	$1/x_j^*$
1	1.1	0.909 091	1	1.05	0.952 381	6	1.55	0.645 161
2	1.3	0.769 231	2	1.15	0.869 565	7	1.65	0.606 061
3	1.5	0.666 667	3	1.25	0.800 000	8	1.75	0.571 429
4	1.7	0.588 235	4	1.35	0.740 741	9	1.85	0.540 541
5	1.9	0.526 316	5	1.45	0.689 655	10	1.95	0.512 821
Sum $S = 3.459\ 540$			Sum $S_1 = 4.052\ 342$			Sum $S_2 = 2.876\ 013$		
$\ln 2 \approx \frac{1}{5}S = 0.691\ 91$			$\ln 2 \approx \frac{1}{10}(S_1 + S_2) = 0.692\ 84$			Error: 0.001 24 (0.2%)		
Computation for $n = 5$			Computation for $n = 10$			Error: 0.000 31 (0.05%)		

We see that the result for $n = 10$ is more accurate than that for $n = 5$.

We shall now derive another integration formula. We subdivide the interval $a \leq x \leq b$ into n equal parts of length $\Delta x = (b - a)/n$, as before, and denote the end points of these subintervals by $a, x_1, x_2, \dots, x_{n-1}, b$.

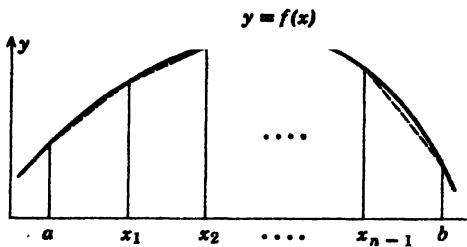


Fig. 29. Trapezoidal rule.

Then the n trapezoids in Fig. 29 have the areas

$$\frac{1}{2}[f(a) + f(x_1)] \Delta x, \quad \frac{1}{2}[f(x_1) + f(x_2)] \Delta x, \dots, \quad \frac{1}{2}[f(x_{n-1}) + f(b)] \Delta x.$$

Thus

$$(3) \quad \int_a^b f(x) dx \approx \Delta x \left[\frac{f(a)}{2} + f(x_1) + f(x_2) + \dots + f(x_{n-1}) + \frac{f(b)}{2} \right]$$

where $\Delta x = (b - a)/n$. This formula is called the **trapezoidal rule**.

Let A be the exact value of the integral under consideration, and let A^* be the approximate value obtained by the trapezoidal rule (3). Then the difference

$$E_T = A^* - A$$

is called the *error* of A^* , and

$$(4) \quad A = \int_a^b f(x) dx = A^* - E_T.$$

Clearly, if $f(x)$ is a *linear* function, its graph is a straight line, and the trapezoidal rule will yield the exact value of the integral, that is, $E_T = 0$. Now in this case, f' is constant and f'' is zero for all x . Hence the error results from the fact that in general $f'' \neq 0$, and it seems to be plausible that its magnitude will depend on the values of f'' in the interval of integration and also on the number n of subintervals in (3). In fact, it can be shown that E_T always lies in the interval

$$(5) \quad KM^* \leq E_T \leq KM$$

where M^* and M are the smallest and largest values of the second derivative of f in the interval of integration, and

$$K = \frac{(b - a)^3}{12n^2}.$$

Example 2. Compute the integral

$$I = \int_0^1 e^{-x^2} dx$$

by means of the trapezoidal rule (3), taking $n = 10$, and estimate the error.

j	x_j	x_j^2	$e^{-x_j^2}$	
0	0	0	1.000 000	
1	0.1	0.01		0.990 050
2	0.2	0.04		0.960 789
3	0.3	0.09		0.913 931
4	0.4	0.16		0.852 144
5	0.5	0.25		0.778 801
6	0.6	0.36		0.697 676
7	0.7	0.49		0.612 626
8	0.8	0.64		0.527 292
9	0.9	0.81		0.444 858
10	1.0	1.00	0.367 879	
		Sums	1.367 879	6.778 167

We have $\Delta x = 0.1$, and

$$I \approx 0.1 \left[\frac{1.363 879}{2} + 6.778 176 \right] = 0.746 211.$$

Estimate of error.

$$f''(x) = 2(2x^2 - 1)e^{-x^2}, \quad M^* = f''(0) = -2, \quad M = f''(1) = 0.735 758.$$

From this and (5),

$$-2K \leq E_T \leq 0.735 758 K$$

where $K = 1/1200$. Thus

$$-0.001 667 \leq E_T \leq 0.000 614.$$

From this and (4) it follows that the exact value of I must lie between

$$0.746 211 - 0.000 614 = 0.745 597$$

and

$$0.746 211 + 0.001 667 = 0.747 878.$$

(The exact value is $I = 0.746 824 \dots$)

The rectangular rule is obtained by approximating the integrand $f(x)$ by a step function (piecewise constant function). The trapezoidal rule results by approximating $f(x)$ by linear functions. We may expect to obtain a more accurate integration formula by approximating the curve of $f(x)$ by portions of parabolas.

For this purpose we subdivide the interval of integration $a \leq x \leq b$ into an even number of equal subintervals, say, into $2n$ subintervals of length $\Delta x = (b - a)/2n$, with end points $x_0 (= a)$, x_1 , \dots , x_{2n-1} , $x_{2n} (= b)$. In the first two intervals we approximate the curve of $f(x)$ by the parabola of the form $\alpha x^2 + \beta x + \gamma$ passing through the points A_0 , A_1 , A_2 of that curve

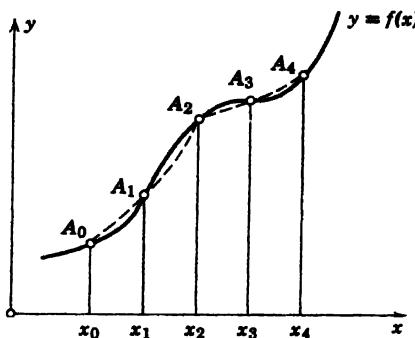


Fig. 30. Simpson's rule.

(Fig. 30). In the next two intervals we approximate that curve by another such parabola through A_2 , A_3 , A_4 , and so on. Proceeding in this fashion, we obtain a curve consisting of n portions of parabolas, and the area under that curve is an approximation for the area under the curve of $f(x)$ between a and b . It can be shown that the integration formula thus obtained is

$$(6) \quad \int_a^b f(x) dx \approx \frac{\Delta x}{3} [f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \cdots + 2f(x_{2n-2}) + 4f(x_{2n-1}) + f(x_{2n})]$$

where $\Delta x = (b - a)/2n$ (cf. Ref. [A13] in Appendix 1). This important formula for approximate integration is called **Simpson's rule**.

If we write (6) in the form

$$(7) \quad \int_a^b f(x) dx = \frac{\Delta x}{3} [f(x_0) + 4f(x_1) + \cdots + f(x_{2n})] - E_S,$$

then E_S is the error of the approximation. It can be shown that E_S lies in the interval

$$(8) \quad CM_4^* \leq E_S \leq CM_4$$

where M_4^* and M_4 are the smallest and the largest value of the fourth derivative of f in the interval of integration, and

$$C = \frac{(b - a)^5}{180(2n)^4}.$$

From (8) we see that Simpson's rule yields the exact result not only for polynomials of the second degree but even for polynomials of the third degree.

Example 3. Evaluate $I = \int_0^1 e^{-x^2} dx$ by Simpson's rule with $2n = 10$ and estimate the error.

j	x_j	x_j^2	$e^{-x_j^2}$		
0	0	0	1.000 000		
1	0.1	0.01		0.990 050	
2	0.2	0.04			0.960 789
3	0.3	0.09		0.913 931	
4	0.4	0.16			0.852 144
5	0.5	0.25		0.778 801	
6	0.6	0.36			0.697 676
7	0.7	0.49		0.612 626	
8	0.8	0.64			0.527 292
9	0.9	0.81		0.444 858	
10	1.0	1.00	0.367 879		
Sums			1.367 879	3.740 266	3.037 901

We find, since $\Delta x = 0.1$,

$$I \approx \frac{0.1}{3} (1.367\ 879 + 4 \cdot 3.740\ 266 + 2 \cdot 3.037\ 901) = 0.746\ 825.$$

Estimate of error. The fourth derivative of the integrand is

$$f^{IV}(x) = 4(4x^4 - 12x^2 + 3)e^{-x^2}.$$

By considering the derivative of f^{IV} we find that the smallest value of f^{IV} in the interval of integration occurs at $x = x^* = 2.5 + 0.5\sqrt{10}$ and the largest value occurs at $x = 0$. Computing the corresponding values of f^{IV} we obtain in (8)

$$M_4^* = f^{IV}(x^*) = -7.359 \dots \quad \text{and} \quad M_4 = f^{IV}(0) = 12.$$

Furthermore, since $2n = 10$ and $b - a = 1$, in (8),

$$C = 1/1\ 800\ 000 = 0.000\ 000\ 55\dots$$

Therefore

$$-0.000\ 004\dots \leq E_S \leq 0.000\ 006\dots$$

This shows that the first four decimals of the above approximation are correct and, by (7),

$$0.746\ 818 < I < 0.746\ 830.$$

The exact value to six decimal places is $I = 0.746\ 824$, and we see that even five decimals or our result are correct. Our present result is much better than that obtained in Ex. 2 by the trapezoidal rule, while the amount of work is almost the same in both cases.

Problems

Review some integration formulas and methods by integrating

1. $\int \frac{dx}{a^2 + x^2}$
2. $\int \frac{dx}{\sqrt{a^2 - x^2}}$
3. $\int e^{ax} \cos bx \, dx$
4. $\int e^{ax} \sin bx \, dx$
5. $\int \ln x \, dx$
6. $\int \cos^3 x \, dx$

7. $\int \sin^2 \omega x \, dx$

8. $\int \tan x \, dx$

9. $\int \frac{dx}{x^2(x^2+1)^2}$

10. Compute $\int_0^1 x^3 \, dx$ by the rectangular rule (2) with $n = 5$. What is the error?

11. Compute the integral in Prob. 10 by the trapezoidal rule (3) with $n = 5$. What error bounds are obtained from (5)? What is the actual error of the result? Why is this result larger than the exact value?

12. Compute the integral in Ex. 2 by using (2) with $n = 5$.

Using the column of $\sin x$ in Table 1 in Sec. 0.1 evaluate $\int_0^1 \frac{\sin x}{x} \, dx$:

13. By the rectangular rule (2) with $n = 5$.

14. By the trapezoidal rule (3) with $n = 5$.

15. By (3) with $n = 10$.

16. By Simpson's rule with $2n = 2$.

17. By Simpson's rule with $2n = 10$.

18. Evaluate $\int_0^1 x^5 \, dx$ by Simpson's rule with $2n = 10$. What error bounds are obtained from (8)? What is the actual error of the result?

19. Find an approximate value of $\ln 2 = \int_1^2 \frac{dx}{x}$ by Simpson's rule with $2n = 4$. Estimate the error by (8).

20. If subintervals of different lengths $\Delta_1 x, \Delta_2 x, \dots, \Delta_n x$ are chosen, show that the trapezoidal rule assumes the form

$$\int_a^b f(x) \, dx \approx \frac{1}{2}[f(a) + f(x_1)] \Delta_1 x + \dots + \frac{1}{2}[f(x_{n-1}) + f(b)] \Delta_n x.$$

Chapter 1

Ordinary Differential Equations of the First Order

Differential equations are of fundamental importance in engineering mathematics. This is due to the fact that many physical laws and relations appear mathematically in the form of a differential equation. We shall consider various physical and geometrical problems which lead to differential equations and the most important standard methods for solving such equations. These methods will in general involve integration.

We shall pay particular attention to the derivation of the differential equations from given physical situations. This transition from the physical problem to a "mathematical model" is of great practical importance and will be illustrated by typical examples.

The present chapter will be devoted to the simplest ordinary differential equations, the so-called equations of the first order. It includes also numerical methods for obtaining approximate solutions of such equations (Sec. 1.13).

Prerequisite for the present chapter: integral calculus.

Sections which may be omitted in a shorter course: 1.8, 1.11–1.13.

References: Appendix 1, Part B.

Answers to problems: Appendix 2.

1.1 Basic Concepts and Ideas

By an **ordinary differential equation** we mean a relation which involves one or several derivatives of an unspecified function y of x with respect to x ; the relation may also involve y itself, given functions of x , and constants.

For example,

$$(1) \quad y' = \cos x,$$

$$(2) \quad y'' + 4y = 0,$$

$$(3) \quad x^2y'''y' + 2e^x y'' = (x^2 + 2)y^2$$

are **ordinary differential equations**.

The term *ordinary* distinguishes it from a *partial differential equation* which involves partial derivatives of an unspecified function of two or more independent variables. For example,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

is a partial differential equation. In the present chapter we shall consider only ordinary differential equations.

An ordinary differential equation is said to be of *order n*, if the *n*th derivative of *y* with respect to *x* is the highest derivative of *y* in that equation.

The notion of the order of a differential equation leads to a useful classification of the equations into equations of first order, second order, etc. Thus (1) is a first-order equation, (2) is of the second order, and (3) is of the third order.

In the present chapter we shall consider first-order equations. Equations of second and higher order will be considered in Chaps. 2-4.

A function

$$(4) \quad y = g(x)$$

is called a **solution** of a given first-order differential equation on some interval, say, $a < x < b$ (perhaps infinite) if it is defined and differentiable throughout the interval and is such that the equation becomes an identity when *y* and *y'* are replaced by *g* and *g'*, respectively.

For example, the function

$$y = g(x) = e^{2x}$$

is a solution of the first-order equation

$$y' = 2y$$

for all *x*, because

$$g' = 2e^{2x},$$

and by inserting *g* and *g'* the equation reduces to the identity

$$2e^{2x} = 2e^{2x}.$$

Sometimes a solution of a given differential will appear as an implicit function, that is, in the form

$$G(x, y) = 0,$$

and is then called an *implicit solution*, in contrast to the *explicit solution* (4).

For example,

$$x^2 + y^2 = 1$$

$$(y \neq 0)$$

is an implicit solution of the differential equation

$$y' = -\frac{x}{y},$$

as the student may verify. In this case we may easily write the solution in explicit form, but in other cases this may be difficult or even impossible.

The principal task of the theory of differential equations is to find all the solutions of a given differential equation and investigate their properties. We shall consider various standard methods which were developed for that purpose.

A differential equation may have many solutions. Let us illustrate this fact by the following examples.

Example 1. Each of the functions

$$y = \sin x, \quad y = \sin x + 3, \quad y = \sin x - \frac{4}{5}$$

is a solution of the equation (1),

$$y' = \cos x,$$

and we know from calculus that any solution of the equation is of the form

$$(5) \quad y = \sin x + c$$

where c is a constant. If we regard c as arbitrary, then (5) represents the totality of all solutions of the equation (cf. Fig. 31).

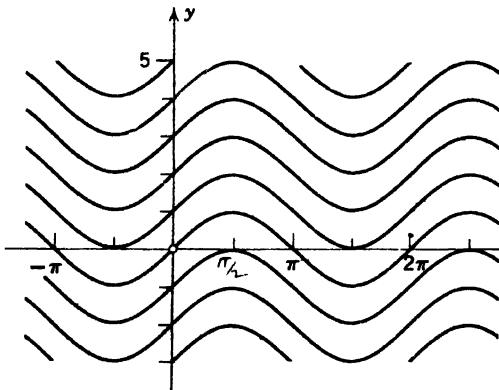


Fig. 31. Solutions of $y' = \cos x$.

Example 2. The student may verify that each of the functions

$$(6) \quad y = e^x, \quad y = 2e^x, \quad y = -\frac{8}{5}e^x$$

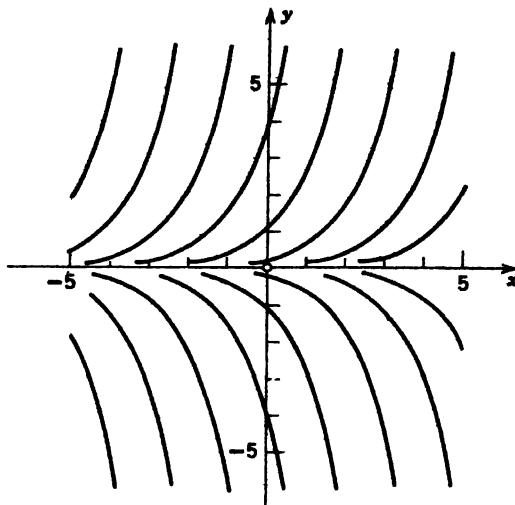
is a solution of the equation

$$y' = y.$$

We shall see later that any solution of the equation is of the form

$$(7) \quad y = ce^x$$

where c is a constant. Regarding c as arbitrary (7) thus represents the totality of all solutions of the equation (cf. Fig. 32 on the next page).

Fig. 32. Solutions of $y' = y$.

Our examples illustrate that a differential equation may (and, in general, will) have more than one solution, even infinitely many solutions, which can be represented by a single formula involving an arbitrary constant c . It is customary to call such a function which contains an arbitrary¹ constant a general solution of the corresponding differential equation of the first order. If we assign a definite value to that constant, then the solution so obtained is called a particular solution.

Thus (7) is a general solution of the equation $y' = y$, and (6) are particular solutions.

In the following sections we shall develop various methods for obtaining general solutions of first order equations. We shall see that each such method will yield a general solution of an equation which is unique, except for notation, and will then be called *the general solution* of that equation.

We mention that in some cases there may be further solutions of a given equation which cannot be obtained by assigning a definite value to the arbitrary constant in the general solution; such a solution is then called a singular solution of the equation.

For example, the equation

$$(8) \quad y'' - xy' + y = 0$$

has the general solution

$$y = cx - c^2$$

representing a family of straight lines where each line corresponds to a definite value of c . A further solution is

$$y = \frac{x^2}{4}$$

¹ The range of the constant may have to be restricted in some cases to avoid imaginary expressions or other degeneracies.

and since this solution cannot be obtained by assigning a definite value to c in the general solution, it is a singular solution. Obviously, each particular solution represents a tangent to the parabola represented by the singular solution (Fig. 33).

Singular solutions will rarely occur in engineering problems.

It should be noted that in some books the notion *general solution* means a formula which includes *all* solutions of an equation, that is, both the particular and the singular ones. We do not adopt this definition for two reasons. First of all, it is frequently quite difficult to prove that a particular formula includes all solutions, and that definition of a general solution is therefore rather useless from the practical point of view. Furthermore, we shall see that a large and very important class of equations (the so-called linear differential equations) does not have singular solutions, and our definition of a general solution can be easily generalized to equations of higher order so that the resulting notion includes all solutions of a differential equation which is linear.

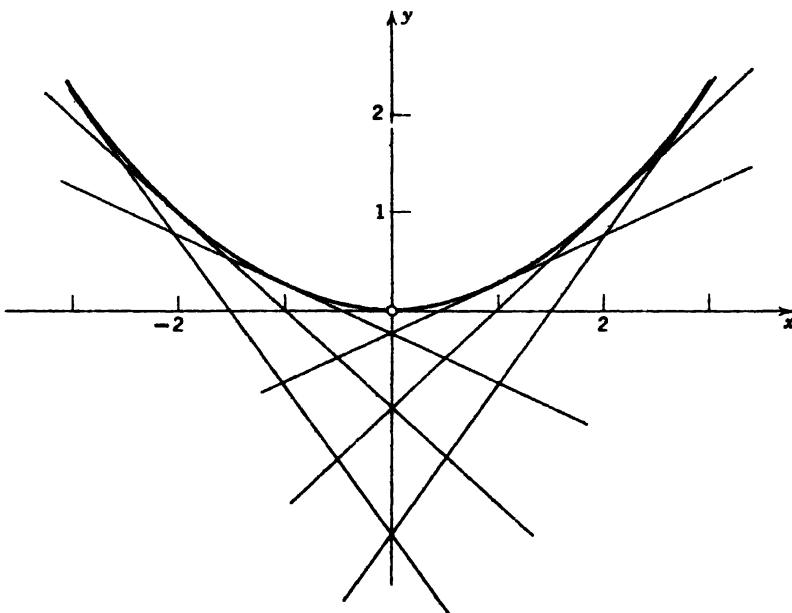


Fig. 33. Singular solution and particular solutions of the equation (8).

We shall see that the conditions under which a given differential equation has solutions are fairly general. But we should note that there are simple equations which do not have solutions at all, and others which do not have a general solution. For example, the equation $y'^2 = -1$ does not have a solution for real y , as is obvious; the equation $|y'| + |y| = 0$ does not have a general solution, because its only solution is $y \equiv 0$.

Differential equations are of great importance in engineering, because many physical laws and relations appear mathematically in the form of differential equations.

Let us consider a simple physical example which will illustrate the typical steps leading from the physical situation to its mathematical formulation

and solution. This may be the easiest way to obtain a first idea of the nature and purpose of differential equations and their applications.

Example 3 (Radioactivity). Experiments show that a radioactive substance decomposes at a rate proportional to the amount present. Starting with a given amount of substance, say, 2 grams, at a certain time, say, $t = 0$, what can be said about the amount available at a later time?

1st Step (Mathematical description of the physical process by a differential equation). We denote the amount of substance still present at time t by $y(t)$. The rate of change is dy/dt . According to the law governing the process of radiation, dy/dt is proportional to y . Thus,

$$(9) \quad -\frac{dy}{dt} = ky$$

where k is a definite physical constant whose numerical value is known for various radioactive substances. (For example, in the case of radium, $k \approx -1.4 \cdot 10^{-11} \text{ sec}^{-1}$.) Clearly, since the amount of substance is positive and decreases with time, dy/dt is negative, and so is k . We see that the physical process under consideration is described mathematically by an ordinary linear differential equation of the first order. Whenever a physical law involves a rate of change of a function, such as velocity, acceleration, etc., it will lead to a differential equation. For this reason differential equations occur frequently in physical and engineering problems.

2nd Step (Solving the differential equation). At this early stage of our consideration no systematic method for solving (9) is at our disposal. However, we see that if (9) has a solution $y(t)$, its derivative must be proportional to y . From calculus we remember that exponential functions have this property. In fact, the function e^{kt} or more generally

$$(10) \quad y(t) = ce^{kt}$$

where c is any constant, is a solution of (9) for all t , as can readily be verified by substituting (10) into (9). Since (10) involves an arbitrary constant, it is a general solution of the first order equation (9). [We shall see that (10) includes all solutions of (9), that is, (9) does not have singular solutions.]

3rd Step (Determination of a particular solution). It is clear that our physical process has a unique behavior, and we may therefore expect that by using further given information we shall be able to select a definite numerical value of c in (10) so that the resulting particular solution will describe the process properly. The amount of substance $y(t)$ still present at time t will depend on the initial amount of substance given. This amount is 2 grams at $t = 0$. Hence we have to specify the value of c so that $y = 2$ when $t = 0$. By inserting this *initial condition*

$$(11) \quad y(0) = 2$$

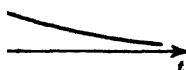
in (10) we obtain

$$y(0) = ce^0 = 2 \quad \text{or} \quad c = 2.$$

If we use this value of c , then (10) becomes

$$(12) \quad y(t) = 2e^{kt}.$$

A y



This particular solution of (9) characterizes the amount of substance still present at any time $t \geq 0$. The physical constant k is negative, and $y(t)$ decreases, as is shown in Fig. 34.

4th Step (Checking). From (12) we have

$$\frac{dy}{dt} = 2ke^{kt} = ky \quad \text{and} \quad y(0) = 2e^0 = 2.$$

Fig. 34. Radioactivity (exponential decay).

The function (12) satisfies the equation (9) as well as the initial condition (11).

The student should never forget to carry out the important final step which shows him whether the function obtained is (or is not) the solution of the problem.

Let us illustrate that geometrical problems may also lead to differential equations.

Example 4. Find the curve through the point $(1, 1)$ in the xy -plane having at each of its points the slope $-y/x$. Clearly the function representing the desired curve must be a solution of the differential equation

$$(13) \quad y' = -\frac{y}{x}.$$

We shall soon learn how to solve such an equation. For the time being the student may verify that

$$y = \frac{c}{x}$$

is a solution of (13) for any value of the constant c . Some of the corresponding curves are shown in Fig. 35. Now we must have $y = 1$ when $x = 1$, and this yields $c = 1$. Hence the solution of our problem is the hyperbola

$$y = \frac{1}{x}.$$

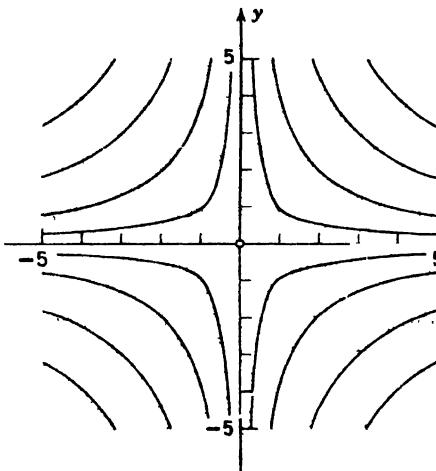


Fig. 35. Solutions of $y' = -y/x$.

Problems

In each case state the order of the differential equation and verify that the given function is a solution.

1. $y' + 2y = 0, \quad y = ce^{-2x}$
2. $y''' = 0, \quad y = ax^3 + bx + c$
3. $y'' + y = 0, \quad y = A \cos x + B \sin x$
4. $y'' - y = x, \quad y = c_1 e^x + c_2 e^{-x} - x$

In each case verify that the given function is a solution of the corresponding differential equation and graph the corresponding curves for some values of the constant c .

5. $y' = 2x$, $y = x^2 + c$ 6. $y' = 2y/x$, $y = cx^2$

7. $y' + 2xy = 0$, $y = ce^{-x^2}$ ("bell-shaped curves")

8. $y' = -x/y$, $x^2 + y^2 = c$ 9. $y' - y = e^{2x}$, $y = ce^{2x} + e^{2x}$

10. $y'^2 - xy' + y = 0$, $y = cx - c^2$, $y = x^2/4$

In each case find a first-order differential equation involving both y and y' for which the given function y is a solution.

11. $y = 4x^3$

12. $y = 1/x$

13. $y = \tan x$

14. $y = 3e^{-4x}$

In each case find $y = \int f(x) dx$ and determine the constant of integration c such that y satisfies the given condition.

15. $f = x^2$, $y = 0$ when $x = 2$

16. $f = \cos^2 x$, $y = \pi/2$ when $x = \pi$

17. $f = \cos 2x$, $y = 1$ when $x = 0$

18. $f = xe^{-x^2}$, $y = 0$ when $x = 0$

In each case verify that the given function is a solution of the corresponding differential equation and determine c such that the resulting particular solution satisfies the given condition. Graph this solution.

19. $y' + y = 0$, $y = ce^{-x}$, $y = 3$ when $x = 0$

20. $y' + y = 5$, $y = ce^{-x} + 5$, $y = 6$ when $x = 1$

21. $y' + 2xy = 0$, $y = ce^{-x^2}$, $y = -2$ when $x = 0$

22. Find the curve in the xy -plane passing through the point $(3, 2)$ and having at each of its points the slope x^2 . Graph the curve.

23. Consider Ex. 3 and find the solution $y(t)$ corresponding to the initial amount 0.01 gram.

24. (Atmospheric pressure). Observations show that the rate of change of the atmospheric pressure p with altitude h is proportional to the pressure. Assuming that the pressure at 18,000 ft is $\frac{1}{2}$ of its value p_0 at sea level, find the formula for the pressure at any height.

25. (Falling body). If a body falls in vacuum due to the action of gravity, starting at time $t = 0$ with initial velocity 0, experiments show that its velocity is proportional to the time. State this physical law in terms of a first-order differential equation, and show that by solving it one obtains the familiar law

$$s(t) = \frac{g}{2} t^2$$

where s is the distance of the body from its initial position and g is the constant of proportionality between velocity and time.

1.2 Geometrical Considerations. Isoclines

We shall now start with a systematic treatment of first-order differential equations. Any such equation may be written in the *implicit form*

(1) $F(x, y, y') = 0.$

Not always, but in many cases we shall be able to write a first-order equation in the *explicit form*

$$(2) \quad y' = f(x, y).$$

Before we consider various standard methods for solving equations of the form (2) we shall demonstrate that (2) has a very simple geometric interpretation. This will immediately lead to a useful graphical method for obtaining a rough picture of the particular solutions of (2), without solving the equation.

We suppose that $f(x, y)$ is *single-valued*, that is, at each point (x, y) of the xy -plane where $f(x, y)$ is defined, it has just one value. The solutions of (2) can be plotted as curves in the xy -plane. We do not know the solutions, but we see from (2) that a solution passing through a point (x_0, y_0) must have the slope $f(x_0, y_0)$ at this point. This suggests the following method. We first graph some of the curves in the xy -plane along which $f(x, y)$ is constant. These curves are called curves of constant slope, or isoclines of (2). Along any such isocline, given by

$$f(x, y) = k = \text{const}$$

we draw a number of parallel short line segments (*lineal elements*) with slope k , which is the slope of solution curves of (2) at any point of that isocline. This we do for all isoclines which we graphed before. In this way we obtain a field of lineal elements, called the *direction field* of (2). With the help of the lineal elements we can now easily graph approximation curves to the (unknown) solution curves of the given equation (2) and thus obtain a qualitatively correct picture of these solution curves.

Let us consider two simple illustrative examples.

Example 1. Graph the direction field of the equation

$$(3) \quad y' = -\frac{x}{y}$$

and an approximation to the solution curve through the point $(\sqrt{5}, 2)$. The isoclines are the straight lines

$$-\frac{x}{y} = k = \text{const} \quad \text{or} \quad y = -\frac{x}{k}.$$

Since the slope of any such line is $-1/k$, the lineal elements are perpendicular to the corresponding isocline. For example, along the isocline $y = x$ we have $x/y = 1$ and from (3), $y' = -1$, etc. (see Fig. 36).

In the next section we shall see that (3) can be easily solved, so that our example merely illustrates the technical details of the method of the direction field, and the same is true for the following example.

Example 2. Graph the direction field of the equation

$$(4) \quad y' = xy$$

and an approximation to the solution curve through the point $(1, 1)$. The isoclines are the equilateral hyperbolas $xy = k$ together with the two coordinate axes. The result is shown in Fig. 37.

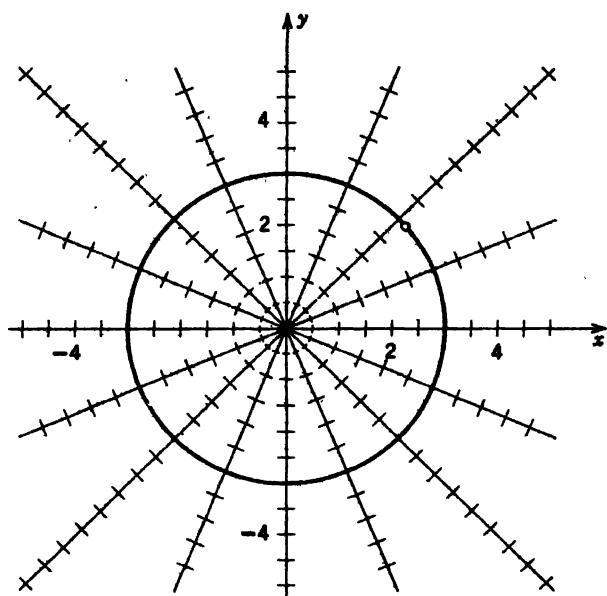


Fig. 36. Direction field of the equation (3).

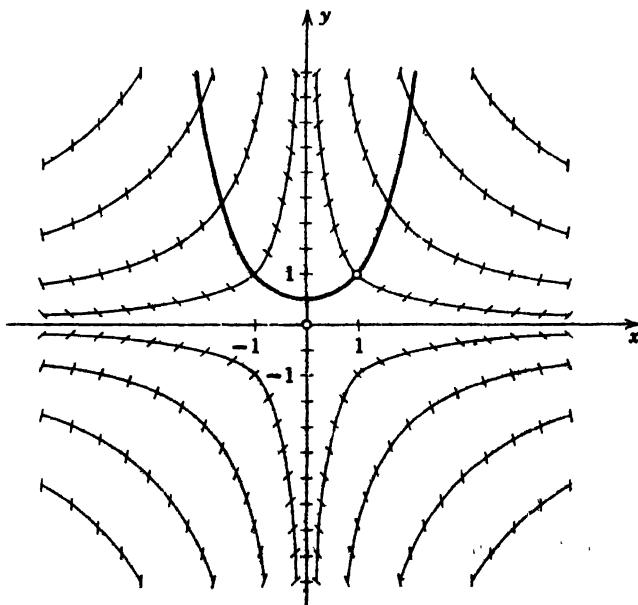


Fig. 37. Direction field of the equation (4).

The method of the direction field is particularly helpful in engineering problems involving a first-order equation whose solution cannot be expressed in terms of known functions or is a very complicated expression.

Problems

- Verify that $x^2 + y^2 = c$ and $y = ce^{x^2/2}$ are the general solutions of (3) and (4) respectively. Compare the exact and the approximate solution curves obtained before.
- Draw the direction field of the differential equation $y' + 2xy = 0$. Plot several approximate solution curves, in particular the curve through the point $(0, 1)$. Verify that the general solution of the equation is $y = ce^{-x^2}$. Compare the exact and the approximate solution curves.

Draw good direction fields for the following differential equations and plot several approximate solution curves.

3. $y' = x$	4. $y' = -2x$	5. $y' + y = 0$
6. $y' = \sin x$	7. $y' = \sin y$	8. $y' = y/x$
9. $y' = 2x + 2y$	10. $y' = x^2 - y$	11. $4yy' + x = 0$.

Using the method of direction fields, plot an approximate solution curve of the given differential equation satisfying the given condition.

- $y' = 2x$, $y = 4$ when $x = 2$
- $y' + 2y = 0$, $y = -3$ when $x = 0$
- $y' + 4x^3y = 0$, $y = 5$ when $x = 0$

Give examples of differential equations (2) whose isoclines are

- Parallel straight lines
- Congruent quadratic parabolas with the y -axis as common axis
- Concentric circles
- What are the isoclines of (2), if f depends on one of the two variables only?
- (Approximate integration). How can we use direction fields for approximate integration of definite integrals? Using this method, evaluate $\int_0^1 x^2 dx$ and $\int_0^1 e^{-x^2} dx$.
- A body B moves on a straight line L , starting at $t = 0$ from the point $s = 1$ where $s(t)$ is the distance of B measured from a fixed point O of L . Suppose that the velocity of B equals $-s'(t)$. Find the corresponding differential equation. Graph the direction field and plot the (approximate) solution curve of the problem.

1.3 Separable Equations

Many first-order differential equations can be reduced by algebraic manipulations to the form

$$(1) \quad g(y)y' = f(x).$$

Since $y' = dy/dx$ we find it convenient to write (1) in the form

$$(2) \quad g(y) dy = f(x) dx,$$

but we keep in mind that this is merely another way of writing (1). Such an equation is called an *equation with separable variables*, or a *separable equation*, because the variables x and y have been *separated* from each other in such a way that x appears only on the right and y appears only on the left. By integrating on both sides of (2) we obtain

$$(3) \quad \int g(y) dy = \int f(x) dx + c.$$

Assuming that f and g are continuous functions, the integrals in (3) will exist, and by evaluating these integrals we obtain the general solution of (1).

Example 1. Solve

$$y' = -\frac{4x}{9y}.$$

By separating variables we have

$$y dy = -\frac{4}{9} x dx.$$

By integrating on both sides we obtain the general solution

$$\frac{y^2}{2} = -\frac{2}{9} x^2 + \tilde{c} \quad \text{or} \quad \frac{y^2}{4} + \frac{x^2}{9} = \frac{\tilde{c}}{2}.$$

The solution represents a family of ellipses.

Example 2. Find the solutions of

$$y' = -2xy$$

in the upper half plane ($y > 0$). By separating variables we have

$$\frac{dy}{y} = -2x dx.$$

Integration yields

$$(4) \quad \ln y = -x^2 + \tilde{c}.$$

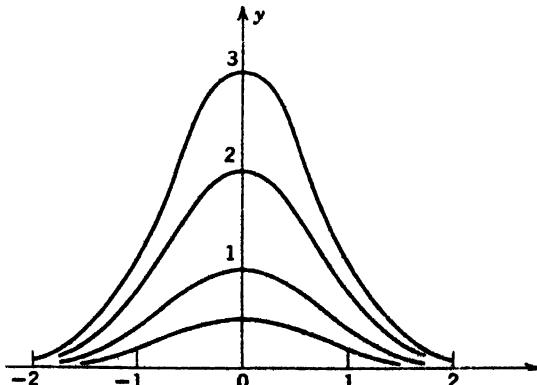


Fig. 38. Solutions of $y' = -2xy$ ("bell-shaped curves").

It is of great importance that the constant of integration is introduced immediately when the integration is carried out. From (4) we obtain

$$y = e^{-x^2 + \tilde{c}}$$

or by noting that $e^{a+b} = e^a e^b$ and setting $e^{\tilde{c}} = c$,

$$y = ce^{-x^2}.$$

The solution represents a family of so-called *bell-shaped curves* (Fig. 38).

Example 3. Solve the differential equation

$$y' = 1 + y^2.$$

By separating variables and integrating we obtain

$$\frac{dy}{1+y^2} = dx, \quad \text{arc tan } y = x + c, \quad y = \tan(x + c).$$

In many engineering applications we are not interested in the general solution of a given differential equation but only in the particular solution $y(x)$ satisfying a given **initial condition**, say, the condition that at some point x_0 the solution $y(x)$ has a prescribed value y_0 , briefly

$$(5) \quad y(x_0) = y_0.$$

- ✓ A first-order differential equation together with an initial condition is called an **initial value problem**. To solve such a problem we must find the particular solution of the equation satisfying the given initial condition.

Example 4. Solve the initial value problem

$$(x^2 + 1)y' + y^2 + 1 = 0, \quad y(0) = 1.$$

1st Step. Separating variables, we find

$$\frac{dy}{1+y^2} = -\frac{dx}{1+x^2}.$$

By integration,

$$\text{arc tan } y = -\text{arc tan } x + c$$

or

$$\text{arc tan } y + \text{arc tan } x = c.$$

Taking the tangent on both sides, we have

$$(6) \quad \tan(\text{arc tan } y + \text{arc tan } x) = \tan c.$$

Now the addition formula for the tangent is

$$\tan(a + b) = \frac{\tan a + \tan b}{1 - \tan a \tan b}.$$

For $a = \text{arc tan } y$ and $b = \text{arc tan } x$ this becomes

$$\tan(\text{arc tan } y + \text{arc tan } x) = \frac{y + x}{1 - xy}.$$

Consequently, (6) may be written

$$(7) \quad \frac{y + x}{1 - xy} = \tan c.$$

2nd Step. We determine c from the initial condition. Setting $x = 0$ and $y = 1$ in (7), we have $1 = \tan c$, so that

$$\frac{y+x}{1-xy} = 1 \quad \text{or} \quad y = \frac{1-x}{1+x}.$$

3rd Step. Check the result.

Example 5 (Newton's law of cooling). A copper ball is heated to a temperature of 100°C . Then at time $t = 0$ it is placed in water which is maintained at a temperature of 30°C . At the end of 3 minutes the temperature of the ball is reduced to 70°C . Find the time at which the temperature of the ball is reduced to 31°C .

Physical information. Experiments show that the rate of change of the temperature T of the ball is proportional to the difference between T and the temperature of the surrounding medium (*Newton's law of cooling*).

Heat flows so rapidly in copper that at any time the temperature is practically the same at all points of the ball.

1st Step. The mathematical formulation of Newton's law of cooling in our case is

$$(8) \quad \frac{dT}{dt} = -k(T - 30)$$

where we denoted the constant of proportionality by $-k$ in order that $k > 0$.

2nd Step. The general solution of (8) is obtained by separating variables; we find

$$T(t) = ce^{-kt} + 30.$$

3rd Step. The given initial condition is $T(0) = 100$. The particular solution satisfying this condition is

$$T(t) = 70e^{-kt} + 30.$$

4th Step. k can be determined from the given information $T(3) = 70$. Thus

$$T(3) = 70e^{-3k} + 30 = 70 \quad \text{or} \quad k = \frac{1}{3} \ln \frac{7}{4} = 0.1865.$$

Hence the temperature $T(t)$ of the ball is

$$T(t) = 70 e^{-0.1865t} + 30,$$

and the value $T = 31^{\circ}\text{C}$ is reached when

$$70 e^{-0.1865t} = 1 \quad \text{or} \quad 0.1865t = \ln 70, \quad t = \frac{\ln 70}{0.1865} = 22.78,$$

that is, after approximately 23 minutes.

5th Step. Check the result.

Example 6 (Velocity of escape from the earth). Find the minimum initial velocity of a body which is fired in radial direction from the earth and is supposed to escape from the earth. Neglect the air resistance and the gravitational pull of other celestial bodies.

1st Step. According to *Newton's law of gravitation*, the gravitational force, and therefore also the acceleration a of the body, is proportional to $1/r^2$ where r is the distance between the body and the center of the earth. Thus

$$a(r) = \frac{dv}{dt} = \frac{k}{r^2}.$$

where v is the velocity and t is the time. Since v is decreasing, $a < 0$ and thus $k < 0$. Let R be the radius of the earth. When $r = R$ then $a = -g$, the acceleration of gravity at the surface of the earth. Note that the minus sign occurs because g is positive and the attraction of gravity acts in the negative direction (the direction towards the center of the earth). Thus

$$-g = a(R) = \frac{k}{R^2} \quad \text{and} \quad a(r) = -\frac{gR^2}{r^3}.$$

Now $v = dr/dt$ and

$$a = \frac{dv}{dt} = \frac{dv}{dr} \frac{dr}{dt} = \frac{dv}{dr} v.$$

Hence the differential equation for the velocity is

$$\frac{dv}{dr} v = -\frac{gR^2}{r^2}.$$

2nd Step. Separating variables and integrating, we obtain

$$(9) \quad v dv = -gR^2 \frac{dr}{r^2} \quad \text{and} \quad \frac{v^2}{2} = \frac{gR^2}{r} + c.$$

3rd Step. On the earth's surface, $r = R$ and $v = v_0$, the initial velocity. For these values of r and v , (9) becomes

$$\frac{v_0^2}{2} = \frac{gR^2}{R} + c, \quad \text{and} \quad c = \frac{v_0^2}{2} - gR.$$

By inserting this expression for c into (9) we obtain

$$(10) \quad v^2 = \frac{2gR^2}{r} + v_0^2 - 2gR.$$

4th Step. If $v^2 = 0$, then $v = 0$, the body stops, and it is clear from the physical situation that then the velocity will change from positive to negative and the body will return to the earth. Consequently, we have to choose v_0 so large that this cannot happen. If we choose

$$(11) \quad v_0 = \sqrt{2gR},$$

then in (10) the expression $v_0^2 - 2gR$ is zero and v^2 is never zero for any r . However, if we choose a smaller value for the initial velocity v_0 , then $v = 0$ for a certain r . The expression in (11) is called the *velocity of escape* from the earth. Since $R = 3960$ statute miles and $g = 32.17 \text{ ft/sec}^2 = 0.00609 \text{ miles/sec}^2$,

$$v_0 = \sqrt{2gR} \approx 7 \text{ miles/sec.}$$

5th Step. Check the result.

Example 7. (Torricelli's law²). A funnel (Fig. 39), whose angle at the outlet is 60° and whose outlet has a cross-sectional area of 0.5 cm^2 , contains water. At time $t = 0$ the outlet is opened and the water flows out. Determine the time when the funnel will be empty, assuming that the initial height of water is $h(0) = 10 \text{ cm}$.

1st Step. The volume of water which flows out during a short interval of time Δt is

$$\Delta V = 0.5 v \Delta t$$

² EVANGELISTA TORRICELLI (1608-1647), Italian physicist. The "contraction factor" 0.6 was introduced by BORDA in 1766 because of the fact that the cross section of the outflowing stream of liquid is somewhat smaller than that of the orifice.

where v is the velocity of the outflowing water. From Torricelli's law, the velocity with which a liquid issues from an orifice is

$$v = 0.6\sqrt{2gh}$$

where $g = 980$ (cm/sec²) is the acceleration of gravity at the surface of the earth and h is the instantaneous height of the liquid above the orifice. Hence

$$\Delta V = 0.3\sqrt{2gh} \Delta t.$$

ΔV must equal the change ΔV^* of the volume of water in the funnel. Now

$$\Delta V^* = -\pi r^2 \Delta h$$

where Δh is the decrease of the height $h(t)$ of the water and $r = h \tan 30^\circ = h/\sqrt{3}$ is the radius of the funnel at height $h(t)$, so that

$$\Delta V^* = -\pi \frac{h^2}{3} \Delta h.$$

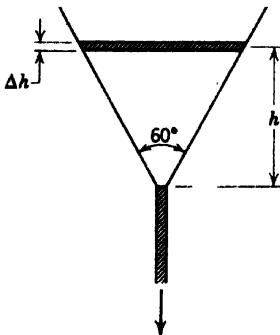


Fig. 39. Funnel.

The minus sign appears because the volume of water in the funnel decreases.

By equating ΔV and ΔV^* we obtain

$$0.3\sqrt{2gh} \Delta t = -\frac{\pi}{3} h^2 \Delta h \quad \text{or} \quad \frac{\Delta h}{\Delta t} = -\frac{0.9\sqrt{2g}}{\pi} h^{-\frac{5}{2}}.$$

If we now let Δt approach zero, we obtain the differential equation

$$\frac{dh}{dt} = -kh^{-\frac{5}{2}}$$

where

$$k = \frac{0.9\sqrt{2g}}{\pi} = 12.7.$$

2nd Step. By separating variables and integrating we have

$$(12) \quad h^{\frac{3}{2}} dh = -k dt \quad \text{and} \quad \frac{2}{5}h^{\frac{5}{2}} = -kt + c.$$

3rd Step. The initial condition is $h(0) = 10$. Setting $t = 0$ and $h = 10$ in the general solution given by the last formula in (12), we obtain

$$\frac{2}{5}10^{\frac{5}{2}} = c.$$

If we use this value of c , the last formula in (12) becomes

$$(13) \quad \frac{2}{5}h^{\frac{5}{2}} = -kt + \frac{2}{5}10^{\frac{5}{2}}.$$

4th Step. By solving (13) for t and substituting $k = 12.7$ we find

$$t = \frac{2}{5k} (10^{\frac{5}{2}} - h^{\frac{5}{2}}) \approx 10 - 0.0315 h^{\frac{5}{2}}.$$

Hence the funnel will be empty ($h = 0$) at $t \approx 10$ sec.

Problems

1. Using (3), show that the solution of (1) satisfying (5) is obtained from

$$(14) \quad \int_{y_0}^y g(y^*) dy^* = \int_{x_0}^x f(x^*) dx^*.$$

2. Consider Ex. 4 and find the solution satisfying $y(2) = 0$.
 3. Why is it important that the constant of integration is introduced immediately when the integration is carried out?

Find the general solution of the following equations (where a , b , ω , and g are constants).

4. $y' + y = 0$

5. $y' = ay/x$

6. $y' = x^2/y$

7. $axy' = by \quad (a \neq 0)$

8. $y' = 1 + x + y^2 + xy^2$

9. $yy' = \cos^2 \omega x$

10. $xyy' = 2(y + 1)$

11. $y^2y' + x^2 = 0$

12. $y' + ay + b = 0 \quad (a \neq 0)$

13. $y' + 2y \tanh x = 0$

14. $y' = y \tan 2x$

15. $xy' = 2\sqrt{y - 1}$

16. $y' + \csc y = 0$

17. $y' \sin 2x = y \cos 2x$

18. $(x^2 + x + 1)y' + (6x + 3)y = 0$

19. $v \frac{dv}{dt} = g$

20. $\frac{dr}{d\theta} = ar$

Solve the following initial value problems.

21. $y' + ky = 0, \quad y(0) = 3.5$

22. $xy' = 3y, \quad y(2) = -8$

23. $y' = y \tanh x, \quad y(0) = y_0$

24. $y' = 4x^3e^{-y}, \quad y(1) = 0$

25. $xy' + y = 0, \quad y(1) = 1$

26. $y' + 3xy = 0, \quad y(0) = \pi$

27. $y' = y \tan x, \quad y(\pi) = -1$

28. $(x + 1)y' = y, \quad y(-2) = -3$

29. $y'x \ln x = 2y, \quad y(2) = (\ln 2)^2$

30. $(y + 2)y' = \sin x, \quad y(0) = 0$

31. $xyy' = y + 3, \quad y(1) = 0$

32. $y^2y' = \cos^2 x, \quad y(0) = 0$

✓ 33. $2yy' = \sin^2 \omega x, \quad y(0) = 1$

34. $y' = y^2 \sin x, \quad y(2\pi) = 1$

35. $\frac{dr}{dt} = -2rt, \quad r(0) = r_0$

36. $v \frac{dv}{dt} = g, \quad v(t_0) = v_0$

37. (Newton's law of cooling) Consider Ex. 5. Find the time at which the temperature of the ball is reduced to 31°C , assuming that at the end of 5 minutes the temperature of the ball is reduced to 60°C while all the other conditions remain as in Ex. 5.

38. (Velocity of escape) Consider the following modification of Ex. 6. The body is carried by a rocket and is separated from the rocket at a distance of 200 miles from the earth's surface. Determine the minimum velocity at this point sufficient for escape from the earth.

39. (Torricelli's law) Consider a spherical tank of radius $R = 50$ cm which contains water and has an outlet of radius $r_0 = 5$ cm at the bottom. At time $t = 0$ the outlet is opened and the water flows out. Determine the time when the tank will be empty, assuming that the initial height of water is $h(0) = R = 50$ cm.

40. In Prob. 39, determine the time necessary to let 50% of the water flow out. Why is this time less than half of the time necessary to drain the tank completely?

41. (Half-life) Experiments show that radium disintegrates at a rate proportional to the amount of radium instantaneously present. Its *half-life*, that is, the time in which 50% of a given amount will disappear, is 1590 years. What per cent will disappear in 1 year?

42. Determine the time in which 1% of each of the following radioactive substances disappears.

Name	Symbol	Isotope of	Discovered in	Half-life
Radium A	$^{216}_{88}\text{Po}$	polonium	1905 (Rutherford)	3.05 minutes
Radium B	$^{214}_{82}\text{Pb}$	lead	1905 (Rutherford)	26.8 minutes
Radium C	$^{214}_{83}\text{Bi}$	bismuth	1905 (Rutherford)	19.7 minutes
Radium D	$^{210}_{82}\text{Pb}$	lead	1903 (P. and M. Curie)	22 years

43. Suppose that 2 g of a radioactive substance is present at time $t = 1$ and 1 g at $t = 2$. Find the initial amount.
44. (Exponential growth) Suppose that in a culture of yeast at each instant the time rate of change of active ferment $y(t)$ is proportional to the amount present. If $y(t)$ doubles in 2 hours, how much can be expected at the end of 8 hours at the same rate of growth?
45. (Exponential decay) Experiments show that the rate of inversion of cane sugar in dilute solution is proportional to the concentration $y(t)$ of the unaltered sugar. Suppose that the concentration is 1/100 at $t = 0$ and is 1/300 at $t = 10$ [hours]. Find $y(t)$.
46. (Evaporation) A wet porous substance in the open air loses its moisture at a rate proportional to the moisture content. If a sheet hung in the wind loses half its moisture during the first hour, when will it have lost 99%, weather conditions remaining the same?
47. Lambert's³ law of absorption states that the absorption of light in a very thin transparent layer is proportional to the thickness of the layer and to the amount incident on that layer. Formulate this in terms of a differential equation and solve it.
48. The law of mass action states that, if the temperature is kept constant, the velocity of a chemical reaction is proportional to the product of the concentrations of the substances which are reacting. In the bimolecular reaction



a moles per liter of a substance A and b moles of a substance B are combined. If y is the number of moles per liter which have reacted after time t , the rate of reaction is given by

$$y' = k(a - y)(b - y).$$

Solve this equation, assuming that $a \neq b$.

49. A tank contains 100 gal of brine in which 10 lb of salt is dissolved. Fresh water runs into the tank at the rate of 5 gal/min, and the mixture, kept practically uniform by stirring, runs out at the same rate. How much salt will there be in the tank at the end of 2 hours?
50. Two liquids are boiling in a vessel. It is found that the ratio of the quantities of each passing off as vapor at any instant is proportional to the ratio of the quantities x and y still in the liquid state. Show that $\frac{dy}{dx} = k \frac{y}{x}$ and solve this equation.
51. Find the curve through the point (3, 4) having at each of its points the slope $-x/y$.
52. Find the curve through the point (2, 0) having at each of its points the slope $-x/4y$.

³ JOHANN HEINRICH LAMBERT (1728–1777), German physicist and mathematician.

53. Find the curves for which the segment of every tangent between the coordinate axes is bisected by the point of tangency.
54. Find the curves for which the slope of the normals at all points equals the ratio of the abscissa to the ordinate.
55. Find the curves whose normals all pass through the origin.

1.4 Equations Reducible to Separable Form

Certain first-order differential equations are not separable but can be made separable by a simple change of variables. This holds for equations of the form⁴

$$(1) \quad y' = g\left(\frac{y}{x}\right)$$

where g is any given function of y/x , for example $(y/x)^3$, $\sin(y/x)$, etc. The form of the equation suggests that we set

$$\frac{y}{x} = u,$$

remembering that y and u are functions of x . Then $y = ux$, and by differentiation,

$$(2) \quad y' = u + u'x.$$

By inserting this into (1) we obtain

$$u + u'x = g(u).$$

Now we may separate variables, finding

$$\frac{du}{g(u) - u} = \frac{dx}{x}.$$

Integrating and replacing u by y/x gives us the general solution of (1).

Example 1. Solve

$$2xyy' - y^2 + x^2 = 0.$$

Dividing by x^2 , we have

$$2\frac{y}{x}y' - \left(\frac{y}{x}\right)^2 + 1 = 0.$$

⁴ These equations are sometimes called homogeneous equations. We shall not use this terminology but reserve the term "homogeneous" for a much more important purpose (cf. Sec. 1.7).

If we set $u = y/x$ and use (2), the equation becomes

$$2u(u + u'x) - u^2 + 1 = 0 \quad \text{or} \quad 2xuu' + u^2 + 1 = 0.$$

Separating variables, we find

$$\frac{2u \, du}{1 + u^2} = -\frac{dx}{x}$$

and by integration

$$\ln(1 + u^2) = -\ln|x| + c^* \quad \text{or} \quad 1 + u^2 = \frac{c}{x}$$

Replacing u by y/x , we finally obtain

$$x^2 + y^2 = cx \quad \text{or} \quad \left(x - \frac{c}{2}\right)^2 + y^2 = \frac{c^2}{4}.$$

Sometimes the form of a given differential equation suggests other simple substitutions, as the following example illustrates.

Example 2. Solve

$$(2x - 4y + 5)y' + x - 2y + 3 = 0.$$

We set $x - 2y = v$. Then $y' = \frac{1}{2}(1 - v')$ and the equation becomes

$$(2v + 5)v' = 4v + 11.$$

Separating variables and integrating, we find

$$\left(1 - \frac{1}{4v + 11}\right)dv = 2 \, dx \quad \text{and} \quad v - \frac{1}{4} \ln|4v + 11| = 2x + c^*$$

or, since $v = x - 2y$,

$$4x + 8y + \ln|4x - 8y + 11| = c.$$

Further simple substitutions are illustrated by the equations in Probs. 11–16.

Problems

Find the general solutions of the following equations.

1. $xy' = x + y$
2. $xy' = xe^{-y/x} + y$
3. $xy' + y + 2x = 0$
4. $(x^2 + 1)y(xy' - y) = x^3$
5. $x^2y' + xy = x^2 + y^2$
6. $2xyy' = 3y^2 + x^2$
7. $y' = \frac{y}{x} + x \sec \frac{y}{x}$
8. $xy' = y + x^2 \tan \frac{y}{x}$
9. $xy' - y - (y - x)^3 = 0$
10. $(xy' - y) \cos(y/x) + x = 0$
11. $(x + 1)(y' - 1) = 2(y - x)$ (put $y - x = v$)
12. $y' = 1 + (y - x) \cot x$ (put $y - x = v$)
13. $y' = \frac{y - x}{y + x}$
14. $y' = \frac{y - x}{y - x + 2}$
15. $y' = \cot(y + x) - 1$
16. $xy' = e^{-xy} - y$ ($xy = v$)
17. Four flies sit at the corners of a card table, facing inward. They start walking simultaneously at the same rate, each directing its motion steadily toward the fly on its right. Find the path of each.

1.5 • Exact Differential Equations

A first-order differential equation

$$(1) \quad M(x, y) dx + N(x, y) dy = 0$$

is said to be **exact** if the left-hand side is the total or exact differential

$$(2) \quad du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy$$

of some function $u(x, y)$. Then (1) may be written

$$du = 0,$$

and by integration we immediately obtain the general solution of (1) in the form

$$(3) \quad u(x, y) = c.$$

By comparing (1) and (2) we see that (1) is exact, if there is some function $u(x, y)$ such that

$$(4) \quad (a) \quad \frac{\partial u}{\partial x} = M, \quad (b) \quad \frac{\partial u}{\partial y} = N.$$

Suppose that M and N are defined and have continuous first partial derivatives in a region in the xy -plane whose boundary is a closed curve having no self-intersections. Then from (4),

$$\frac{\partial M}{\partial y} = \frac{\partial^2 u}{\partial x \partial y}, \quad \frac{\partial N}{\partial x} = \frac{\partial^2 u}{\partial y \partial x}$$

and by assumption of continuity the second derivatives are equal [cf. (4) in Sec. 0.2]. Thus

$$(5) \quad \frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}.$$

This condition is not only necessary but also sufficient⁵ for $M dx + N dy$ to be a total differential.

If (1) is exact, the function $u(x, y)$ can be found by guessing or in the following systematic way. From (4a) we have⁶

$$(6) \quad u = \int M dx + k(y);$$

⁵ We shall prove this fact at another occasion (in Sec. 6.12); the proof can also be found in some books on elementary calculus; cf. Ref. [A13] in Appendix 1.

⁶ Or from (4b), $u = \int N dy + k(x)$, etc.; see Ex. 1.

in this integration, y is to be regarded as a constant, and $k(y)$ plays the role of a "constant" of integration. To determine $k(y)$, we find $\partial u / \partial y$ from (6), use (4b) to find dk/dy , and integrate.

Example 1. Solve

$$xy' + y + 4 = 0.$$

If we write the equation in the form (!), that is,

$$(y + 4) dx + x dy = 0,$$

we see that $M = y + 4$, $N = x$, and (5) holds, so that the equation is exact. Now from (4b) we have $N = \partial u / \partial y = x$, and by integration

$$u = xy + k(x).$$

To determine $k(x)$ we differentiate u and apply (4a), finding

$$\frac{\partial u}{\partial x} = y + \frac{dk}{dx} = M = y + 4.$$

Thus $dk/dx = 4$ and $k = 4x + c^*$. By inserting this into the equation $u = xy + k = \text{const}$ and solving for y , we obtain the general solution of our equation,

$$y = \frac{c}{x} - 4.$$

Intelligent students, who know that thinking is an important thing in mathematics, write the given equation in the form

$$y dx + x dy = -4 dx,$$

recognize the left side as the total differential of xy , and integrate, finding $xy = -4x + c$, which is equivalent to our previous result.

Example 2. Solve

$$2x \sin 3y dx + 3x^2 \cos 3y dy = 0.$$

The equation is exact, and from (6) we obtain

$$u = \int 2x \sin 3y dx + k(y) = x^2 \sin 3y + k(y).$$

Therefore $\partial u / \partial y = 3x^2 \cos 3y + dk/dy$, and $dk/dy = 0$ or $k = c^* = \text{const}$. The general solution is $u = \text{const}$ or $x^2 \sin 3y = c$.

Problems

Find the total differential of the given functions $u(x, y)$.

- | | |
|---------------------------------|--|
| 1. $u = x^2 + 2xy^2 - \sin(xy)$ | 2. $u = \ln(x^2 + y^2) + 2 \arctan(y/x)$ |
| 3. $u = e^x \cos y + (x - y)^2$ | 4. $u = \tan(x^2y^2) - x^3 + y^3$ |

Find the exact equations that have the following general solutions.

- | | |
|------------------------|--------------------------|
| 5. $x^3y + \sin x = c$ | 6. $x^2 + \cosh(xy) = c$ |
| 7. $xe^{-2xy} = c$ | 8. $e^x \sin y = c$ |

Are the following equations exact? Find the general solution.

9. $4y \, dy + x \, dx = 0$

10. $y \cos(xy) \, dx + x \cos(xy) \, dy = 0$

11. $x \, dy + 2y^2 \, dx = 0$

12. $2 \cosh x \, dx + \sinh x \, dy = 0$

13. $4 \sin y \, dx + x \cos y \, dy = 0$

14. $\sinh x \cos y \, dx = \cosh x \sin y \, dy$

✓15. $\left(3x^2y + \frac{y}{x}\right) \, dx + (x^3 + \ln x) \, dy = 0$

16. $\frac{\cos 2y}{x} \, dx = 2 \ln x \sin 2y \, dy$

✓17. $(xe^{xy} + 2y)y' + ye^{xy} = 0$

18. $x^3y' \sec^2 y + 3x^2 \tan y = \cos x$

Solve the following initial value problems.

19. $9(y - 1) \, dy + (x - 3) \, dx = 0, \quad y(3) = 0$

20. $2xy \, dx + (x^2 + \pi \cos \pi y) \, dy = 0, \quad y(1) = 1$

21. $(\cos x + y \sin x) \, dx = \cos x \, dy, \quad y(\pi) = 0$

22. $xe^{x^2+y^2} \, dx + y(e^{x^2+y^2} + 1) \, dy = 0, \quad y(0) = 0$

Find and graph the curve having the given slope y' and passing through the given point (x, y) .

23. $y' = \frac{1-x}{1+y}, \quad (1, 0)$

24. $y' = \frac{x}{4-4y}, \quad (0, 2)$

25. $y' = \frac{x+2}{y+1}, \quad (-3, -1)$

26. $y' = -\frac{y-2}{x-2}, \quad (0, 0)$

Solve the following equations by integrating by inspection.

27. $x \, dy + y \, dx = 0$

28. $xy^2 \, dx + x^2y \, dy = 0$

29. $\cos x \cos y \, dx - \sin x \sin y \, dy = 0$

30. $2x \, dx + x^{-2}(x \, dy - y \, dx) = 0$

Under what conditions are the following equations exact? Solve the exact equations.

✓31. $(2x + ay) \, dx + (2y + bx) \, dy = 0$

32. $(\cosh y + a \cos ax) \, dx + bx \sinh y \, dy = 0$

Under what conditions are the following equations exact?

33. $f(x)g(y) \, dx + h(y) \, dy = 0$

34. $[f(x) + g(y)] \, dx + [h(x) + l(y)] \, dy = 0$

✓1.6 Integrating Factors

Sometimes a given differential equation

(1) $P(x, y) \, dx + Q(x, y) \, dy = 0$

is not exact but can be made exact by multiplying it by a suitable function $F(x, y)$ ($\neq 0$). This function is then called an **integrating factor** of (1). With some experience, integrating factors can be found by inspection. For this purpose the student should keep in mind the differentials listed in Ex. 2 below. In some important special cases, integrating factors can be determined in a systematic way which is illustrated in the subsequent section.

Example 1. Solve

$$x \, dy - y \, dx = 0.$$

The equation is not exact. An integrating factor is $F = 1/x^2$, and

$$F(x)(x \, dy - y \, dx) = \frac{x \, dy - y \, dx}{x^2} = d\left(\frac{y}{x}\right) = 0, \quad y = cx.$$

Theorem 1. If (1) is not exact and has a general solution $u(x, y) = c$, then there exists an integrating factor (even infinitely many such factors).

Proof. By differentiating $u(x, y) = c$ we have

$$du = \frac{\partial u}{\partial x} \, dx + \frac{\partial u}{\partial y} \, dy = 0.$$

From this and (1) it follows that

$$\frac{\partial u}{\partial x} : \frac{\partial u}{\partial y} = P : Q$$

must hold identically. Hence a function $F(x, y)$ exists such that

$$\frac{\partial u}{\partial x} = FP, \quad \frac{\partial u}{\partial y} = FQ,$$

and then

$$(2) \quad F(x, y)(P \, dx + Q \, dy) = du.$$

This shows that F is an integrating factor of (1). Furthermore, multiplication of (2) by a function $H(u)$ yields the expression

$$H(u) F(P \, dx + Q \, dy) = H(u) \, du$$

which is exact. Hence $H(u)F(x, y)$ is another integrating factor, and since H is arbitrary, there are infinitely many integrating factors of (1).

Example 2. Find integrating factors of $x \, dy - y \, dx = 0$; cf. Ex. 1. Since

$$d\left(\frac{x}{y}\right) = \frac{y \, dx - x \, dy}{y^2}, \quad d\left(\ln \frac{y}{x}\right) = \frac{x \, dy - y \, dx}{xy}, \quad d\left(\arctan \frac{y}{x}\right) = \frac{x \, dy - y \, dx}{x^2 + y^2}$$

the functions $1/y^2$, $1/xy$, and $1/(x^2 + y^2)$ are such factors. The corresponding solutions

$$\frac{x}{y} = c, \quad \ln \frac{y}{x} = c, \quad \arctan \frac{y}{x} = c$$

are essentially the same, because each represents a family of straight lines through the origin.

Problems

Find integrating factors F and solve:

✓ 1. $2y^2 \, dx + 3xy \, dy = 0$

2. $y \, dx + 2x \, dy = 0$

✓ 3. $y \, dx + x \ln x \, dy = 0$

4. $\sin y \, dx + \cos y \, dy = 0$

$$5. y \, dx - 2x \, dy = 0$$

$$7. x \cos y \, dy - \sin y \, dx = 0$$

$$6. (y + 1) \, dx - (x + 1) \, dy = 0$$

$$8. 2 \cos x \cos y \, dx - \sin x \sin y \, dy = 0$$

Find integrating factors F and solve the following initial value problems.

$$9. 2y \, dx + x \, dy = 0, \quad y(2) = -1$$

$$10. 2y \, dx - x \, dy = 0, \quad y(1) = 2$$

$$11. \cos y \, dx = \sin y \, dy, \quad y(0) = \pi$$

$$12. \cosh y \, dx + 2x \sinh y \, dy = 0, \quad y(1) = 0$$

Find conditions under which the following functions F are integrating factors of (1). (Hint: assume $F(P \, dx + Q \, dy) = 0$ to be exact and use (5) in Sec. 1.5.)

$$13. F = y^k$$

$$14. F = x^m$$

$$15. F = e^x$$

$$16. F = e^{xy}$$

$$17. F = x^m y^k$$

18. Using the result of Prob. 14, solve Probs. 3, 7, 9, and 10.

19. Show that if $\frac{1}{Q} \left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x} \right)$ is a function of x only, say, $f(x)$, then $F(x) = e^{\int f(x) \, dx}$ is an integrating factor of (1).

20. Show that Probs. 14 and 15 are special cases of Prob. 19.

1.7 Linear First-Order Differential Equations

A first-order differential equation is said to be **linear** if it can be written in the form

$$(1) \quad y' + f(x)y = r(x).$$

The characteristic feature of this equation is that it is linear in y and y' , whereas f and r may be *any* given functions of x .

If $r(x) \equiv 0$, the equation is said to be **homogeneous**; otherwise, it is said to be **nonhomogeneous**. (Here, $r \equiv 0$ means $r = 0$ for all x .)

Let us find a formula for the general solution of (1) in some interval I , assuming that f and r are continuous in I . For the homogeneous equation

$$(2) \quad y' + f(x)y = 0$$

this is very simple. By separating variables we have

$$\frac{dy}{y} = -f(x) \, dx \quad \text{and thus} \quad \ln |y| = - \int f(x) \, dx + c^*$$

or

$$(3) \quad y(x) = ce^{-\int f(x) \, dx} \quad (c = e^{c^*}).$$

To solve the nonhomogeneous equation (1), let us write it in the form

$$(fy - r) \, dx + dy = 0$$

and show that we can find an integrating factor $F(x)$ depending only on x . If such a factor exists,

$$F(x)(fy - r) \, dx + F(x) \, dy = 0$$

must be exact. Now for this equation the condition (5) of Sec. 1.5 assumes the form

$$\frac{\partial}{\partial y} [F(fy - r)] = \frac{dF}{dx} \quad \text{or} \quad Ff = \frac{dF}{dx}$$

By separating variables and integrating we obtain

$$\ln F = \int f(x) dx$$

and from this

$$F(x) = e^{h(x)} \quad \text{where} \quad h(x) = \int f(x) dx.$$

This shows that $F(x)$ is an integrating factor of (1). Let us now multiply (1) by this factor, finding

$$e^h(y' + fy) = e^h r.$$

Since $h' = f$, this may be written

$$\frac{d}{dx}(ye^h) = e^h r.$$

Now we can integrate on both sides and obtain

$$ye^h = \int e^h r dx + c.$$

By dividing both sides by e^h we have the desired formula

$$(4) \quad y(x) = e^{-h} \left[\int e^h r dx + c \right], \quad h = \int f(x) dx,$$

which represents the general solution of (1) in the form of an integral. (The choice of the values of the constants of integration is immaterial; cf. Prob. 2 below.)

Example 1. Solve the linear differential equation

$$y' - y = e^{2x}.$$

Here

$$f = -1, \quad r = e^{2x}, \quad h = \int f dx = -x$$

and from (4) we obtain the general solution

$$y(x) = e^{-x} \left[\int e^{-x} e^{2x} dx + c \right] = e^{-x}[c e^x + c] = ce^x + e^{2x}.$$

Example 2. Solve

$$xy' + y + 4 = 0.$$

We write this equation in the form (1), that is,

$$y' + \frac{1}{x}y = -\frac{4}{x}.$$

Hence, $f = \frac{1}{x}$, $r = -\frac{4}{x}$ and, therefore,

$$h = \int f dx = \ln x, \quad e^h = x, \quad e^{-h} = \frac{1}{x}.$$

From this and (4) we obtain the general solution

$$y(x) = \frac{1}{x} \left[\int x \left(-\frac{4}{x} \right) dx + c \right] = \frac{c}{x} - 4,$$

in agreement with Ex. 1 of Sec. 1.5.

Of course, in simple cases such as Exs. 2, 3, and 5 (below) we may solve the equation without using (4).

Example 3. Solve the linear equation

$$xy' + y = \sin x.$$

We can write the equation in the form

$$d(xy) = \sin x dx$$

and integrate on both sides, finding

$$xy = -\cos x + c \quad \text{or} \quad y = \frac{1}{x}(c - \cos x).$$

The procedure of solving initial value problems may be illustrated by the following examples.

Example 4. Solve the initial value problem

$$y' + y \tan x = \sin 2x, \quad y(0) = 1.$$

Here $f = \tan x$, $r = \sin 2x = 2 \sin x \cos x$, and

$$\int f dx = \int \tan x dx = \ln |\sec x|.$$

From this we see that in (4),

$$e^h = \sec x, \quad e^{-h} = \cos x, \quad e^h r = 2 \sin x$$

and the general solution of our equation is

$$y(x) = \cos x \left[2 \int \sin x dx + c \right] = c \cos x - 2 \cos^2 x.$$

According to the initial condition, $y = 1$ when $x = 0$, that is

$$1 = c - 2 \quad \text{or} \quad c = 3,$$

and the solution of our initial value problem is

$$y = 3 \cos x - 2 \cos^2 x.$$

Example 5. Solve the initial value problem

$$x^2 y' + 2xy - x + 1 = 0, \quad y(1) = 0.$$

The equation may be written as $d(x^2y) = (x - 1) dx$. By integration

$$x^2y = \frac{1}{2}x^2 - x + c \quad \text{or} \quad y(x) = \frac{1}{2} - \frac{1}{x} + \frac{c}{x^2}.$$

From this and the initial condition, $y(1) = \frac{1}{2} - 1 + c = 0$ or $c = \frac{1}{2}$. Thus

$$y(x) = \frac{1}{2} - \frac{1}{x} + \frac{1}{2x^2}.$$

Very often the solutions of differential equations appearing in engineering applications are not elementary functions. In the case of equation (1) they appear as integrals, as can be seen from (4). If it happens that such a non-elementary integral is a tabulated function,⁷ that is a great convenience, but

it is not vital. The essential thing is to reduce the solution to a computable form, for example, by approximate integration, such as Simpson's rule, or by developing the integrand in a power series and integrating term by term.



Fig. 40. Simple block diagram.

In applications the independent variable x will often be the time; the function $r(x)$ on the right-hand side of (1) may represent a force, and the solution $y(x)$ a displacement, a current, or some other variable physical quantity. In engineering mathematics $r(x)$ is frequently called the **input**, and $y(x)$ is called the **output** or *response to the input* (and the initial conditions). This terminology is widely used in electrical engineering where the differential equation may govern the behavior of an electric circuit and the output $y(x)$ is obtained as the solution of that equation corresponding to the input $r(x)$. For example, in the field of analogue computers the physical system may be an **integrating unit**; then the corresponding differential equation is

$$y' = r(x),$$

and the output $y(x)$ corresponding to an input $r(x)$ is the integral of $r(x)$,

$$y(x) = \int_{x_0}^x r(x^*) dx^*.$$

The system may be represented by a "block diagram" as shown in Fig. 40. Various examples will occur in our further consideration.

Problems

1. Show that $e^{-\ln x} = 1/x$ (but not $-x$) and $e^{-\ln(\sec x)} = \cos x$.
2. Show that the choice of the value of the constant of integration in $\int f dx$ [cf. (4)] is immaterial.

⁷ Whether tables of a certain function were computed can be found in the extremely useful book Ref. [A6], listed in Appendix 1. Some important nonelementary integrals will be considered in Chap. 17, power series in Secs. 3.2 and 13.5, and numerical methods for solving first-order differential equations in Sec. 1.13. Methods of approximate integration are included in Sec. 0.8.

Find the general solutions of the following differential equations.

- | | | |
|----------------------------|---------------------------|--------------------------|
| 3. $y' - y = 1 - x$ | 4. $y' + y = 1$ | 5. $y' - 4y = 2x - 4x^2$ |
| 6. $y' - y = e^x$ | 7. $y' = 2x(y - x) + 1$ | 8. $y' + 2y = \cos x$ |
| 9. $y' - 3y = \sin 2x$ | 10. $y' + 2xy = x$ | 11. $xy' + y = x^3 + x$ |
| 12. $xy' = 2y + x^4$ | 13. $xy' + 2y = 4x^2$ | 14. $xy' + y = 2x$ |
| 15. $y' + 3y = e^{2x} + 6$ | 16. $y' = (y - 1) \cot x$ | 17. $y' - 3y = e^{3x}$ |

Solve the following initial value problems.

- | | |
|------------------------------------|--|
| 18. $y' + 2y = 3, \quad y(0) = 0$ | 19. $y' - 4y = 8x^2, \quad y(0) = -0.25$ |
| 20. $xy' + y = 2x, \quad y(1) = 2$ | 21. $y' + 5y = 3e^x, \quad y(0) = 1$ |

Prove that the linear differential equations (1) and (2) have the following properties.

22. $y \equiv 0$ is a solution of (2), but not of (1) [with $r(x) \neq 0$].
23. If y_1 and y_2 are solutions of (2), then $y_1 + y_2$ is a solution of (2).
24. If y_1 is a solution of (2), then $y = cy_1$ (c any constant) is a solution of (2).
25. If y_1 and y_2 are solutions of (1), then $y = y_1 - y_2$ is a solution of (2).
26. If y_1 is a solution of (1) and y_2 is a solution of (2), then $y = y_1 + y_2$ is a solution of (1).
27. If y_1 and y_2 are solutions of (1), then $y = y_1 + y_2$ satisfies $y' + fy = 2r$.
28. If y_1 is a solution of (1), then $y = cy_1$ satisfies $y' + fy = cr$.
29. The transformation $x = \phi(t)$ (where ϕ has a continuous derivative) transforms (1) into a linear differential equation involving y and dy/dt .
30. The linear transformation $y = a(x)u + b(x)$ with differentiable $a \neq 0$ and b transforms (1) into a linear differential equation involving u and du/dx .

1.8 Variation of Parameters

There is another interesting way of obtaining the general solution of the linear differential equation

$$(1) \quad y' + f(x)y = r(x).$$

We have seen that

$$(2) \quad v(x) = e^{-\int f(x)dx}$$

is a solution of the corresponding homogeneous linear equation. Let us try to determine a function $u(x)$ such that

$$(3) \quad y(x) = u(x)v(x)$$

is the general solution of (1). This attempt is suggested by the form of the general solution $cv(x)$ of the homogeneous equation and consists in replacing the parameter c by a variable $u(x)$. Therefore, this approach, due to Lagrange,⁸

⁸ JOSEPH LOUIS LAGRANGE (1736–1813), a great French mathematician, spent 20 years of his life in Prussia and then returned to Paris. His important major work was in the calculus of variations, celestial and general mechanics, differential equations, and algebra.

is called the **method of variation of parameters**. It can be generalized to equations of higher order where it is of great importance. Our present simple case is a good occasion for getting acquainted with this basic method.

By substituting (3) into (1) we obtain

$$u'v + u(v' + fv) = r.$$

Since v is a solution of the homogeneous equation, this reduces to

$$u'v = r \quad \text{or} \quad u' = \frac{r}{v}.$$

Integration yields

$$u = \int \frac{r}{v} dx + c.$$

We thus obtain the result

$$(4) \quad y = uv = v \left(\int \frac{r}{v} dx + c \right),$$

and from (2) we see that this is identical with (4) of the previous section.

Problems

Apply the method of variation of parameters to the following equations.

1. $y' + y = 2$

2. $(x + 4)y' + 3y = 3$

3. $y' - y = 3e^x$

Certain nonlinear first-order differential equations may be reduced to linear form by a suitable change in the dependent variable, as the following problems illustrate.

4. The equation

$$y' + f(x)y = g(x)y^a \quad (a \text{ any real number})$$

is called **Bernoulli's⁹ equation**. For $a = 1$, the equation is linear, and for $a \neq 1$ it is nonlinear. Set $[y(x)]^{1-a} = u(x)$ and show that the equation assumes the linear form

$$u' + (1 - a)f(x)u = (1 - a)g(x).$$

Solve the following Bernoulli equations.

5. $xy' = y - y^{-1}$

6. $xy' + y = y^2$

7. $3y' + xy = xy^{-2}$

8. $y' + y = x/y$

9. Solve $(y^4 + 2x)y' = y$ by taking x as the dependent variable.

Find substitutions which reduce the following differential equations to linear form and solve the resulting equations.

10. $(y' - 1)e^y = 1 - x$

11. $y' \cos y + \sin y = 5$

12. $y' + xy \ln y = 2xy$

13. $y' \sinh y = \cosh y + e^x$

⁹ JACOB BERNOULLI (1654–1705), Swiss mathematician, who contributed to the theory of elasticity and to the theory of mathematical probability. The method for solving Bernoulli's equation was discovered by Leibniz in 1696.

14. Show that the direction field of (1) has the following interesting property: The linear elements on any vertical line $x = x_0$ all pass through a single point with coordinates

$$\xi = x_0 + \frac{1}{f(x_0)}, \quad \eta = \frac{r(x_0)}{f(x_0)}.$$

How can this property be used for constructing the direction field of a linear differential equation?

15. Applying the result in Prob. 14, find an approximate solution curve of the initial value problem $y' = 2xy + 1$, $y(1) = 1$.



1.9 Electric Circuits

Linear first-order differential equations have various applications in physics and engineering practice. To illustrate this, let us consider some standard examples in connection with electric circuits. These and similar considerations are important, because they help all students to learn how to express physical situations in terms of mathematical relations. The transition from the physical situation to a corresponding mathematical model is always the first step in engineering mathematics, and this important step requires experience and training which can only be gained by considering typical examples from various fields.

The simplest electric circuit is a series circuit in which we have a source of electric energy (*electromotive force*) such as a generator or a battery, and a resistor, which uses energy, for example an electric light bulb (Fig. 41). If we close the switch, a current I will flow through the resistor, and this will cause a *voltage drop*, that is, the electric potential at the two ends of the resistor will be different; this potential difference or voltage drop can be measured by a voltmeter. Experiments show that the following law holds.

The voltage drop E_R across a resistor is proportional to the instantaneous current I , say

$$(1) \quad E_R = RI \quad (\text{Ohm's law})$$

where the constant of proportionality R is called the **resistance** of the resistor. The current I is measured in *amperes*, the resistance R in *ohms*, and the voltage E_R in *volts*.¹⁰

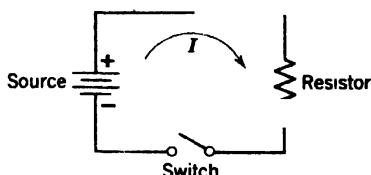


Fig. 41. Circuit.

¹⁰ These and the subsequent units are named after ANDRÉ MARIE AMPÈRE (1775–1836), French physicist, GEORG SIMON OHM (1789–1854), German physicist, ALLESSANDRO VOLTA (1745–1827), Italian physicist, JOSEPH HENRY (1797–1878), American physicist, MICHAEL FARADAY (1791–1867), English physicist, and CHARLES AUGUSTIN DE COULOMB (1736–1806), French physicist.

The other two important elements in more-complicated circuits are *inductors* and *capacitors*. An inductor opposes a change in current, having an inertia effect in electricity similar to that of mass in mechanics; we shall consider this analogy later (Sec. 2.14). Experiments yield the following law.

The voltage drop E_L across an inductor is proportional to the instantaneous time rate of change of the current I , say,

$$(2) \quad E_L = L \frac{dI}{dt}$$

where the constant of proportionality L is called the **inductance** of the inductor and is measured in *henrys*; the time t is measured in seconds.

A capacitor is an element which stores energy. Experiments yield the following law.

The voltage drop E_C across a capacitor is proportional to the instantaneous electric charge Q on the capacitor, say,

$$(3') \quad E_C = \frac{1}{C} Q$$

where C is called the **capacitance** and is measured in *farads*; the charge Q is measured in *coulombs*. Since

$$(3'') \quad I(t) = \frac{dQ}{dt}$$

this may be written

$$(3) \quad E_C = \frac{1}{C} \int_{t_0}^t I(t^*) dt^*.$$

The current $I(t)$ in a circuit may be determined by solving the equation (or equations) resulting from the application of Kirchhoff's¹¹ second law:

The algebraic sum of all the instantaneous voltage drops around any closed loop is zero, or the voltage impressed on a closed loop is equal to the sum of the voltage drops in the rest of the loop.

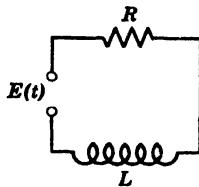
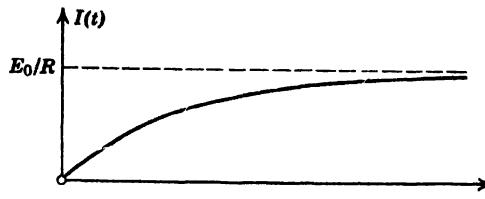
Example 1 (RL-circuit). For the "RL-circuit" in Fig. 42 we obtain from Kirchhoff's second law and (1), (2)

$$(4) \quad L \frac{dI}{dt} + RI = E(t).$$

Case A (Constant electromotive force). If $E = E_0 = \text{const}$, then (4) in Sec. 1.7 yields the general solution

$$(5) \quad \begin{aligned} I(t) &= e^{-\alpha t} \left| \frac{E_0}{L} \int e^{\alpha t} dt + c \right| \\ &= \frac{E_0}{R} + ce^{-(R/L)t}. \end{aligned} \quad (\alpha = R/L)$$

¹¹ GUSTAV ROBERT KIRCHHOFF (1824–1887), German physicist. Later we shall also need Kirchhoff's first law: At any point of a circuit the sum of the inflowing currents is equal to the sum of the outflowing currents.

Fig. 42. *RL*-circuit.Fig. 43. Current in an *RL*-circuit due to a constant electromotive force.

The last term approaches zero as t tends to infinity, so that $I(t)$ tends to the limit value E_0/R ; after a sufficiently long time, I will practically be constant. Figure 43 shows the particular solution

$$(5^*) \quad I(t) = \frac{E_0}{R} (1 - e^{-(R/L)t}),$$

which corresponds to the initial condition $I(0) = 0$.

Case B (Periodic electromotive force). If $E(t) = E_0 \sin \omega t$, then, by (4) in Sec. 1.7, the general solution of (4) in this current section is

$$I(t) = e^{-\alpha t} \left[\frac{E_0}{L} \int e^{\alpha t} \sin \omega t \, dt + c \right] \quad (\alpha = R/L).$$

Integration by parts yields

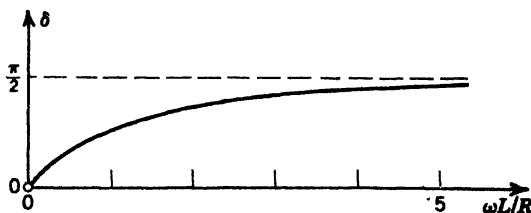
$$I(t) = ce^{-(R/L)t} + \frac{E_0}{R^2 + \omega^2 L^2} (R \sin \omega t - \omega L \cos \omega t).$$

According to (15b) in Sec. 0.1 this may be written

$$(6) \quad I(t) = ce^{-(R/L)t} + \frac{E_0}{\sqrt{R^2 + \omega^2 L^2}} \sin(\omega t - \delta), \quad \delta = \arctan \frac{\omega L}{R}.$$

The exponential term will approach zero as t approaches infinity. This means that after a sufficiently long time the current $I(t)$ executes practically harmonic oscillations. [Cf. Fig. 45.] Figure 44 shows the phase angle δ as a function of $\omega L/R$. If $L = 0$, then $\delta = 0$, and the oscillations of $I(t)$ are in phase with those of $E(t)$.

An electrical (or dynamical) system is said to be in the **steady state** when the variables describing its behavior are periodic functions of the time or constant, and it is said to be in the **transient state** (or *unsteady state*) when it is not in the steady state. The corresponding variables are called **steady state functions** and **transient functions**, respectively.

Fig. 44. Phase angle δ in (6) as a function of $\omega L/R$.

$\uparrow I(t)$

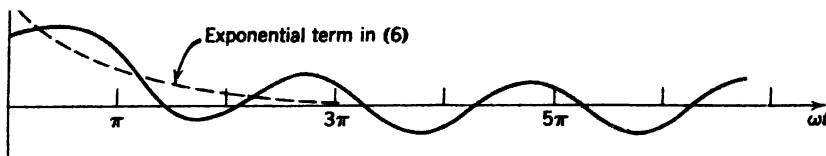


Fig. 45. Current in an RL -circuit due to a sinusoidal electromotive force, as given by (6) (with $\delta = \pi/4$).

In Ex. 1, Case A, the function E_0/R is the steady state function or **steady state solution** of (4), and in Case B the steady state solution is represented by the last term in (6). Before the circuit (practically) reaches the steady state it is in the transient state. It is clear that such an interim or transient period occurs because inductors and capacitors store energy, and the corresponding inductor currents and capacitor voltages cannot be changed instantly. Similar situations arise in various physical systems. For example, if a radio receiver having heater-type vacuum tubes is turned on, a transient interval of time must elapse during which the tubes change from "cold" to "hot."

Example 2 (RC-circuit). By applying Kirchhoff's second law and (1), (3) to the RC -circuit in Fig. 46 we obtain the equation

$$(7) \quad RI + \frac{1}{C} \int I dt = E(t).$$

To get rid of the integral we differentiate with respect to t , finding

$$(8) \quad R \frac{dI}{dt} + \frac{1}{C} I = \frac{dE}{dt}.$$

According to (4) in Sec. 1.7, the general solution of (8) is

$$(9) \quad I(t) = e^{-t/RC} \left(\frac{1}{R} \int e^{t/RC} \frac{dE}{dt} dt + c \right).$$

Case A (Constant electromotive force). If E is constant, then $\frac{dE}{dt} = 0$, and (9) assumes the simple form (Fig. 47)

$$(10) \quad I(t) = ce^{-t/RC}.$$

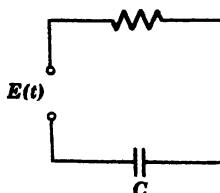


Fig. 46. RC -circuit.



Fig. 47. Current in an RC -circuit due to a constant electromotive force.

Case B (Sinusoidal electromotive force). If $E(t) = E_0 \sin \omega t$, then

$$\frac{dE}{dt} = \omega E_0 \cos \omega t,$$

and by inserting this in (9) and integrating by parts we find

$$(11) \quad I(t) = ce^{-t/RC} + \frac{\omega E_0 C}{1 + (\omega RC)^2} (\cos \omega t + \omega RC \sin \omega t)$$

$$= ce^{-t/RC} + \frac{\omega E_0 C}{\sqrt{1 + (\omega RC)^2}} \sin(\omega t - \delta),$$

where $\tan \delta = -1/\omega RC$. The first term decreases steadily as t increases, and the last term represents the steady-state current, which is sinusoidal. The graph of $I(t)$ is similar to that in Fig. 45.

More complicated circuits and the analogy between electrical and mechanical vibrations will be considered in connection with second-order differential equations in Sec. 2.14.

Problems

1. Derive and check (5).
2. Derive and check (6).
3. Verify (11).
4. In (5*), the current approaches E_0/R . The rapidity of approach depends on R/L . The reciprocal $\tau = L/R$ is called the time constant of the circuit. Graph (5*) with $E_0/R = 1$ and $\tau = 0.1, 1$, and 10 . Express the slope of the tangents to the curves at $t = 0$ in terms of τ .
5. Show that when $t = \tau$ the current (5*) has reached about 63% of its final value.
6. Show that the current (5*) reaches 50% of its final value when $t = \tau \ln 2 \approx 0.7\tau$.
7. Assuming that $L = 10$ henrys, find τ and R such that (5*) reaches 99% of its final value at $t = 1$ sec.
8. Obtain from (5) and graph the particular solutions satisfying the initial conditions $I(0) = 0, E_0/2R, E_0/R$, and $2E_0/R$.
9. In Ex. 1, Case A, let $R = 50$ ohms, $L = 0.2$ millihenry ($2 \cdot 10^{-4}$ henry), and $I(0) = 0$. Find the time when the current reaches 99.9% of its final value.
10. In Ex. 1, Case A, let $R = 10$ ohms, $L = 0.5$ henry, $E_0 = 110$ volts, and $I(0) = 0$. Determine the time constant τ and the time when $I = 10$ amperes.
11. For what initial condition does (6) yield the steady state solution?

Find the current $I(t)$ in the RL -circuit in Fig. 42 under the following assumptions.

12. $R = 10$ ohms, $L = 100$ henrys, $E = 40$ volts when $0 \leq t \leq 100$ sec, $E = 0$ when $t > 100$ sec, $I(0) = 4$
13. $R = 1$ ohm, $L = 100$ henrys, E periodic as shown in Fig. 48, $I(0) = 0$

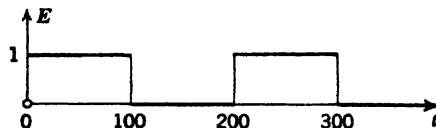


Fig. 48. Problem 13.

14. Show that $I(0) = (e - 1)/(e^2 - 1)$ is the only initial condition for which $I(t)$ in Prob. 13 is periodic, that is, $I(t + 200) = I(t)$ for all $t \geq 0$.
15. For what initial condition does (11) yield the steady state solution?
16. Obtain from (11) the particular solution satisfying $I(0) = 0$.
17. Find the current in the RC -circuit in Fig. 46, assuming that $R = 10$ ohms, $C = 0.1$ farad, $E(t) = 110 \sin 314t$ volts, and $I(0) = 0$.
18. Using (3'), introduce $Q(t)$ as the dependent variable in (7) and solve the resulting equation for the cases of a constant and a sinusoidal electromotive force.
19. A capacitor ($C = 0.01$ farad) in series with a resistor ($R = 20$ ohms) is charged from a battery ($E_0 = 10$ volts), cf. Fig. 46. Find the charge $Q(t)$ and the voltage $V(t)$ on the capacitor, assuming that at $t = 0$ the capacitor is completely uncharged. Find the current $I(t)$ in the circuit.
20. Suppose that the initial charge on the capacitor ($C = 0.1$ farad) in Fig. 49 is 2 coulombs. At $t = 0$ the switch is closed and the capacitor starts discharging. Find the current $I(t)$, the charge $Q(t)$, and the voltage $V(t)$ on the capacitor assuming that $R = 100$ ohms.

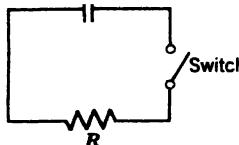


Fig. 49. Problem 20.

21. In the RL -circuit in Fig. 42, let $E = 10$ volts, $R = 5$ ohms, $L = (10 - t)$ henrys when $0 \leq t \leq 10$ sec, $L = 0$ when $t > 10$ sec, and $I(0) = 0$. Find $I(t)$.
22. Solve Prob. 21 when $L = t + 1$ henrys ($0 \leq t \leq 1$ sec), $L = 2$ henrys ($t > 1$ sec), and the other data are as before.
23. In Fig. 42, let $L = 1$ henry, $R = 1/(1 - t)$ ohm when $0 \leq t < 1$, $E = 1$ volt, and $I(0) = 0$. Find $I(t)$ when $0 \leq t < 1$.
24. Find the current $I(t)$ in the RC -circuit (Fig. 46), assuming that $E = 10$ volts, $C = 0.1$ farad, $R = (10 - t)$ ohms when $0 \leq t \leq 10$ sec, $R = 0$ when $t > 10$ sec, and $I(0) = 1$ ampere.
25. Find the charge $Q(t)$ and the voltage $V(t)$ on the capacitor in Prob. 24.

1.10 Families of Curves. Orthogonal Trajectories

If for each fixed real value of c the equation

$$(1) \quad F(x, y, c) = 0$$

represents a curve in the xy -plane and if for variable c it represents infinitely many curves, then the totality of these curves is called a **one-parameter family of curves**, and c is called the *parameter* of the family.

Example 1. The equation

$$(2) \quad F(x, y, c) = x + y + c = 0$$

represents a family of parallel straight lines; each line corresponds to precisely one value of c . The equation

$$(3) \quad F(x, y, c) = x^2 + y^2 - c^2 = 0$$

represents a family of concentric circles of radius c with center at the origin.

The general solution of a first-order differential equation involves a parameter c and thus represents a family of curves. This yields a possibility for representing many one-parameter families of curves by such differential equations. The practical use of such representations will become obvious from our further consideration.

Example 2. By differentiating (2) we see that

$$y' + 1 = 0$$

is the differential equation of that family of straight lines. Similarly the differential equation of the family (3) is

$$y' = -x/y.$$

If the equation obtained by differentiating (1) still contains the parameter c , then we have to eliminate c from this equation by using (1).

Example 3. The differential equation of the family of parabolas

$$(4) \quad y = cx^2$$

is obtained by differentiating (4),

$$(5) \quad y' = 2cx,$$

and by eliminating c from (5). From (4) we have $c = y/x^2$, and by substituting this into (5) we find the desired result

$$(6) \quad y' = 2y/x.$$

In many engineering applications, a family of curves is given, and it is required to find another family whose curves intersect each of the given curves at right angles.¹² Then the curves of the two families are said to be *mutually orthogonal*, they form an *orthogonal net*, and the curves of the family to be obtained are called the **orthogonal trajectories** of the given curves (and conversely); cf. Fig. 50 on the next page.

Let us mention some familiar examples. The meridians on the earth's surface are the orthogonal trajectories of the parallels. On a map the curves of steepest descent are the orthogonal trajectories of the contour lines. In electrostatics the equipotential lines and the lines of electric force are orthogonal trajectories of each other. We shall see later that orthogonal trajectories play an important role in various fields of physics, for example, in hydrodynamics and heat conduction.

¹² Remember that the angle of intersection of two curves is defined to be the angle between the tangents of the curves at the point of intersection.

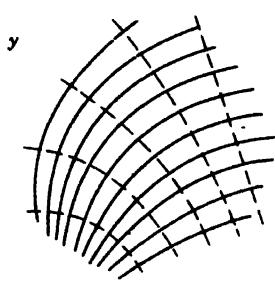


Fig. 50. Curves and their orthogonal trajectories.

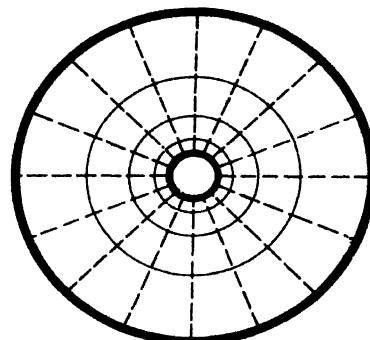


Fig. 51. Equipotential lines and lines of electric force (dashed) between two concentric cylinders.

Given a family $F(x, y, c) = 0$ which can be represented by a differential equation

$$(7) \quad y' = f(x, y)$$

we may find the corresponding orthogonal trajectories as follows. From (7) we see that a curve of the given family which passes through a point (x_0, y_0) has the slope $f(x_0, y_0)$ at this point. The slope of the orthogonal trajectory through (x_0, y_0) at this point should be the negative reciprocal of $f(x_0, y_0)$, that is, $-1/f(x_0, y_0)$, because this is the condition for the tangents of the two curves at (x_0, y_0) to be perpendicular. Consequently, the differential equation of the orthogonal trajectories is

$$(8) \quad y' = -\frac{1}{f(x, y)},$$

and the trajectories are obtained by solving this equation.

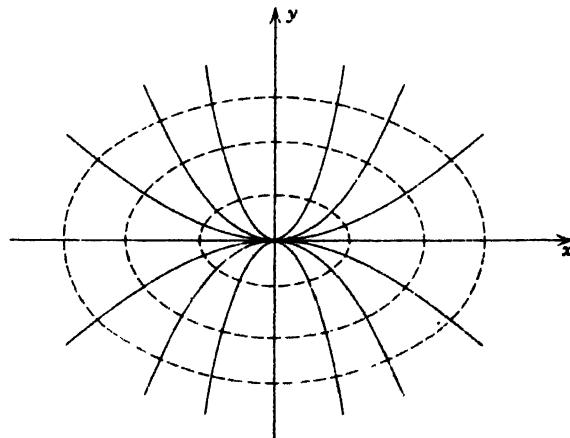


Fig. 52. Parabolas and their orthogonal trajectories in Example 4.

It can be shown that, under relatively general conditions, a family of curves has orthogonal trajectories, but we shall not consider this question here.

Example 4. Find the orthogonal trajectories of the parabolas in Ex. 3. From (6) we see that the differential equation of the orthogonal trajectories is

$$y' = - \frac{1}{2y/x} = - \frac{x}{2y}.$$

By separating variables and integrating we find that the orthogonal trajectories are the ellipses (Fig. 52)

$$\frac{x^2}{2} + y^2 = c^*.$$

Example 5. Find the orthogonal trajectories of the circles

$$(9) \quad x^2 + (y - c)^2 = c^2.$$

We first determine the differential equation of the given family. By differentiating (9) with respect to x we find

$$(10) \quad 2x + 2(y - c)y' = 0.$$

We have to eliminate c . Solving (9) for c , we have

$$c = \frac{x^2 + y^2}{2y}.$$

By inserting this into (10) and simplifying we get

$$x + \frac{y^2 - x^2}{2y} y' = 0 \quad \text{or} \quad y' = \frac{2xy}{x^2 - y^2}.$$

From this and (8) we see that the differential equation of the orthogonal trajectories is

$$y' = - \frac{x^2 - y^2}{2xy} \quad \text{or} \quad 2xyy' - y^2 + x^2 = 0.$$

The orthogonal trajectories obtained by solving this equation (cf. Ex. 1 in Sec. 1.4) are the circles (Fig. 53)

$$(x - \tilde{c})^2 + y^2 = \tilde{c}^2.$$

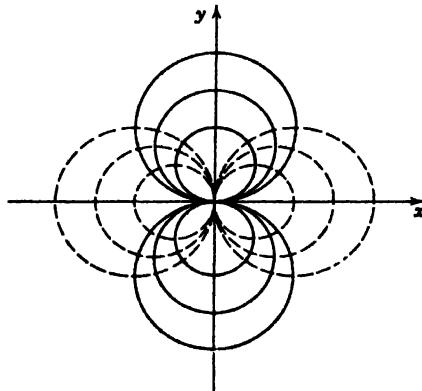


Fig. 53. Circles and their orthogonal trajectories (dashed) in Example 5.

Problems

What families of curves are represented by the following equations?

- | | | |
|---------------------------|----------------------|--------------------------|
| 1. $y - x + c = 0$ | 2. $y - cx = 0$ | 3. $x^2 + (y - c)^2 = 4$ |
| 4. $(x - c)^2 + 4y^2 = 4$ | 5. $x^2 - y^2 = c^2$ | 6. $y - cx^2 = 0$ |
| 7. $y - x^3 = c$ | 8. $y^2 = (x - c)^3$ | 9. $y = (x - c)^3$ |

Represent the following families of curves in the form (1).

10. The straight lines through the point $x = 4, y = -1$
11. The circles which pass through the origin and whose centers lie on the straight line $y = x$
12. The catenaries obtained by translating the catenary $y = \cosh x$ in the direction of the straight line $y = -x$

Represent the following families of curves by differential equations.

- | | | | |
|------------------|--------------------|-------------------|------------------------|
| 13. $xy = c$ | 14. $y = cx$ | 15. $y = ce^{3x}$ | 16. $y = ce^{2x} + 2$ |
| 17. $y = e^{cx}$ | 18. $y = c \sin x$ | 19. $y = cx^3$ | 20. $cx^2 + 4cy^2 = 1$ |

21. Derive (6) by writing (4) in the form $yx^{-2} = c$ and differentiating this equation.

Find the orthogonal trajectories of the following curves. Graph some of the curves and their trajectories.

- | | | |
|-------------------|----------------------|-------------------------|
| 22. $y = 2x + c$ | 23. $y = cx$ | 24. $x^2 + y^2 = c^2$ |
| 25. $y = cx^2$ | 26. $x^2 + 4y^2 = c$ | 27. $xy = c$ |
| 28. $y = ce^{3x}$ | 29. $y = \tan x + c$ | 30. $y = \sqrt{2x + c}$ |

31. Show that the orthogonal trajectories of the family $g(x, y) = c$ can be obtained from the differential equation

$$\frac{dy}{dx} = \frac{\partial g / \partial y}{\partial g / \partial x}.$$

Using the result in Prob. 31, find the orthogonal trajectories of the following curves.

32. $x^2 + 9y^2 = c$

33. $xy = c$

34. $y/x = c$

35. Show that (8) may be written in the form

$$\frac{dx}{dy} = -f(x, y).$$

Using this result, find the orthogonal trajectories of the curves $y = \sqrt{x + c}$.

36. Consider two families of curves $u(x, y) = c$ and $v(x, y) = c^*$. Show that if u and v satisfy the first-order partial differential equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x},$$

then those families are orthogonal. (Those so-called *Cauchy-Riemann differential equations* are basic in complex analysis and will be considered in Chap. 10.)

37. Isogonal trajectories of a family of curves are curves which intersect the members

of the family at a constant angle θ . Show that at each point the slopes m_1 and m_2 of the tangents to the corresponding curves satisfy the relation

$$\frac{m_2 - m_1}{1 + m_1 m_2} = \tan \theta = \text{const.}$$

38. Using the result in Prob. 37, find the curves which intersect the circles $x^2 + y^2 = c$ at an angle of 45° .

1.11 Picard's Iteration Method

We have seen that the solution of a *linear* first-order differential equation can be obtained by evaluating an integral; cf. (4) in Sec. 1.7. In the case of other differential equations of the first order the situation may be more complicated. In fact, there are various such equations which cannot be solved by one of the standard methods considered before or by another elementary method.¹³ In engineering problems that involve such equations it will often be sufficient to obtain approximate solutions. In such cases the method which we shall now consider may be helpful. This method, which is called **Picard's iteration method**,¹⁴ yields an approximate solution of an initial value problem of the form

$$(1) \quad y' = f(x, y), \quad y(x_0) = y_0.$$

Iteration methods are methods which consist of a repeated execution of the same type of steps where in each step the result of the preceding step (or steps) is used; we want to mention that these methods are of increasing importance in various branches of applied mathematics.

The basic idea of Picard's method is very simple. By integration we see that (1) may be written in the form

$$(2) \quad y(x) = y_0 + \int_{x_0}^x f[t, y(t)] dt,$$

where t denotes the variable of integration. In fact, when $x = x_0$ the integral is zero and $y = y_0$, so that (2) satisfies the initial condition in (1); furthermore, by differentiating (2) we obtain the differential equation in (1).

To find approximations to the solution $y(x)$ of (2) we proceed as follows. We substitute the crude approximation $y = y_0 = \text{const}$ on the right; this yields the presumably better approximation

$$y_1(x) = y_0 + \int_{x_0}^x f(t, y_0) dt.$$

¹³ Reference [B6] in Appendix 1 includes more than 1500 important differential equations and their solutions, arranged in systematic order and accompanied by numerous references to original literature.

¹⁴ EMILE PICARD (1856–1941), French mathematician who made important contributions to the theory of complex analytic functions and differential equations.

Then we substitute $y_1(x)$ in the same way to get

$$y_2(x) = y_0 + \int_{x_0}^x f[t, y_1(t)] dt,$$

etc. The n th step of this iteration yields

$$(3) \quad y_n(x) = y_0 + \int_{x_0}^x f[t, y_{n-1}(t)] dt.$$

In this way we obtain a sequence of approximations

$$y_1(x), \quad y_2(x), \quad \dots, \quad y_n(x), \dots.$$

and we shall see in the next section that the conditions under which this sequence converges to the solution $y(x)$ of (1) are relatively general.

Let us now illustrate the practical application of Picard's method by a simple example which can also be treated otherwise so that we can check our results.

Example 1. Find approximate solutions to the initial value problem

$$y' = 1 + y^2, \quad y(0) = 0.$$

In this case, $x_0 = 0$, $y_0 = 0$, $f(x, y) = 1 + y^2$, and (3) becomes

$$y_n(x) = \int_0^x [1 + y_{n-1}^2(t)] dt = x + \int_0^x y_{n-1}^2(t) dt.$$

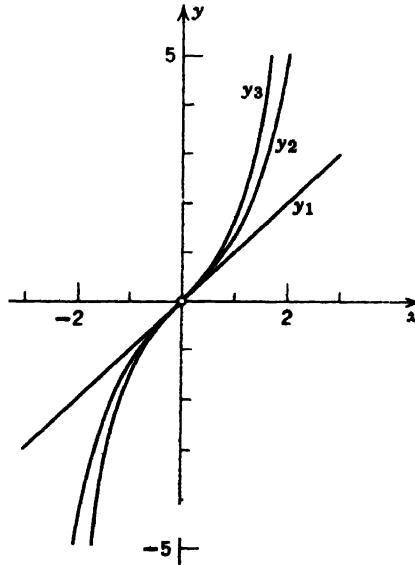


Fig. 54. Approximate solutions in Example 1.

Starting from $y_0 = 0$, we thus obtain

$$\begin{aligned}y_1(x) &= x + \int_0^x 0 \cdot dt = x \\y_2(x) &= x + \int_0^x t^2 dt = x + \frac{1}{3}x^3 \\y_3(x) &= x + \int_0^x \left(t + \frac{t^3}{3}\right)^2 dt = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \frac{1}{945}x^7\end{aligned}$$

etc. Of course, we can obtain the exact solution of our present problem by separating variables (Ex. 3 in Sec. 1.3), finding

$$(4) \quad y(x) = \tan x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \frac{17}{315}x^7 + \dots \quad \left(-\frac{\pi}{2} < x < \frac{\pi}{2}\right).$$

The first three terms of $y_3(x)$ and the series in (4) are the same. The series in (4) converges for $|x| < \pi/2$, and all we may expect is that our sequence y_1, y_2, \dots converges to a function which is the solution of our problem for $|x| < \pi/2$. This illustrates that the problem of convergence is of practical importance.

Problems

1. Apply Picard's iteration method to the problem $y' = y$, $y(0) = 1$ and show that the successive approximations tend to the limit $y = e^x$, the exact solution.
2. Apply Picard's method to $y' = 2xy$, $y(0) = 1$. Graph y_1, y_2, y_3 and the exact solution for $0 \leq x \leq 2$.
3. Compute $y_1(1), y_2(1), y_3(1)$ in Prob. 2 and compare these values with the exact value $y(1) = e = 2.718 \dots$.
4. Apply Picard's method to $y' = 2xy + 1$, $y(0) = 0$ and show that the method yields

$$y_1 = x, \quad y_2 = x + \frac{2}{3}x^3, \quad y_3 = x + \frac{2}{3}x^3 + \frac{4}{15}x^5, \quad \text{etc.}$$

Using (4) in Sec. 1.7, show that these are the first terms of the Maclaurin series of the exact solution

$$y(x) = e^{x^2} \int_0^x e^{-t^2} dt = x + \frac{2}{3}x^3 + \frac{2^2}{3 \cdot 5}x^5 + \frac{2^3}{3 \cdot 5 \cdot 7}x^7 +$$

(The integral in this formula cannot be evaluated by elementary methods of calculus and will be considered later in Sec. 17.2.)

Apply Picard's method to the initial value problems consisting of the differential equation $y' = x + y$ and the initial condition:

$$5. \quad y(0) = 1 \quad 6. \quad y(0) = -1 \quad 7. \quad y(0) = 0$$

Apply Picard's method to the following initial value problems.

- | | |
|-----------------------------------|---------------------------------------|
| 8. $y' = 2y$, $y(0) = -1$ | 9. $y' = xy$, $y(0) = 2$ |
| 10. $y' - 2xy = -1$, $y(0) = 0$ | 11. $y' - xy = 2x - x^3$, $y(0) = 0$ |
| 12. $y' + y^2 = x$, $y(0) = 0.5$ | 13. $y' = xy + 1$, $y(0) = 0$ |
14. Show that both $y \equiv 0$ and $y = -x^2/4$ are solutions of the problem

$$y' = \sqrt{\frac{x^2}{4} + y} - \frac{x}{2}, \quad y(0) = 0,$$

in the range $x \geq 0$. Which of these solutions is obtained by Picard's method, starting from $y_0 = 0$?

15. Show that if f in (1) does not depend on y , the approximations obtained by Picard's method are identical with the exact solution.

1.12 Existence and Uniqueness of Solutions

The initial value problem

$$|y'| + |y| = 0, \quad y(0) = 1$$

has no solution because $y \equiv 0$ is the only solution of the differential equation. The initial value problem

$$y' = x, \quad y(0) = 1$$

has precisely one solution, namely, $y = \frac{1}{2}x^2 + 1$. The initial value problem

$$y' = \frac{y-1}{x}, \quad y(0) = 1$$

has infinitely many solutions, namely, $y = 1 + cx$ where c is arbitrary. From these examples we see that an initial value problem

$$(1) \quad y' = f(x, y), \quad y(x_0) = y_0$$

may have none, precisely one, or more than one solution. This leads to the following two fundamental questions.

Problem of existence. Under what conditions does an initial value problem of the form (1) have at least one solution?

Problem of uniqueness. Under what conditions does that problem have a unique solution, that is, only one solution?

Theorems which state such conditions are called **existence theorems** and **uniqueness theorems**, respectively.

Of course, our three examples are so simple that we can find the answer to those two questions by inspection, without using any theorems. However, it is clear that in more complicated cases—for example, when the equation cannot be solved by elementary methods—existence and uniqueness theorems will be of great importance. As a matter of fact, a more advanced course in differential equations consists mainly of considerations concerning the existence, uniqueness, and general behavior of solutions of various types of differential equations, and many of these considerations have far-reaching practical consequences. For example, results about the existence and uniqueness of periodic solutions of certain differential equations, which were obtained by Poincaré for entirely theoretical reasons about 80 years ago,

are nowadays the base of many practical investigations in nonlinear mechanics.

✓ We want to mention that in many branches of mathematics existence and uniqueness theorems are of similar importance. Let us illustrate this by a familiar example in linear algebra. The three systems of linear equations

$$(a) \begin{array}{l} x + y = 1 \\ x + y = 0 \end{array} \quad (b) \begin{array}{l} x + y = 1 \\ x - y = 0 \end{array} \quad (c) \begin{array}{l} x + y = 1 \\ 2x + 2y = 2 \end{array}$$

have, respectively, none, precisely one, and infinitely many solutions (cf. also Ex. 1 in Sec. 0.3). This is immediately clear, without using any theorems. However, if a system of many equations in many unknowns is given, the need for existence and uniqueness theorems in this field becomes obvious because then the question of whether the system has a solution at all, and if so, how many solutions there are, cannot be answered immediately.

In connection with differential equations, a student who is exclusively interested in applications and does not like theoretical considerations (an attitude which will prevent him from having good success in the applied area) may be inclined to reason in the following manner. The physical problem corresponding to a certain differential equation has a unique solution, and the same must therefore be true for the differential equation. In simpler cases he may be right. However, he should keep in mind that the differential equation is merely an abstraction of the reality obtained by disregarding certain physical facts which seem to be of minor influence, and in complicated physical situations there may be no way of judging a priori the importance of various circumstances. In such a case there will then be no a priori guarantee that the differential equation leads to a faithful picture of the reality.

In the case of an initial value problem of the form

$$(1) \quad y' = f(x, y), \quad y(x_0) = y_0$$

the conditions for the existence and uniqueness of the solution are quite simple. If f is continuous in some region of the xy -plane containing the point (x_0, y_0) , then the problem (1) has at least one solution. If moreover the partial derivative $\partial f / \partial y$ exists and is continuous in that region, then the problem (1) has precisely one solution. This solution can then be obtained by Picard's iteration method. Let us formulate these three statements in a precise manner.

Existence theorem. *If $f(x, y)$ is continuous at all points (x, y) in some rectangle (cf. Fig. 55 on p. 84)*

$$R: \quad |x - x_0| < a, \quad |y - y_0| < b$$

and bounded¹⁶ in R , say,

$$(2) \quad |f(x, y)| \leq K \quad . \quad \text{for all } (x, y) \text{ in } R,$$

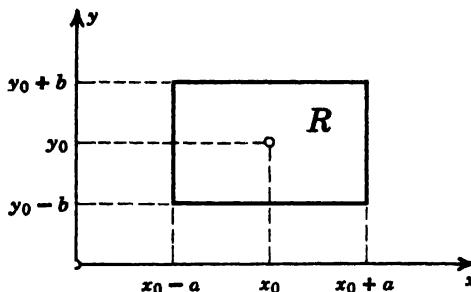


Fig. 55. Rectangle R in the existence and uniqueness theorems.

then the initial value problem (1) has at least one solution $y(x)$, which is defined at least for all x in the interval $|x - x_0| < \alpha$ where α is the smaller of the two numbers a and b/K .

Uniqueness theorem. If $f(x, y)$ and $\partial f / \partial y$ are continuous for all (x, y) in that rectangle R and bounded, say,

$$(3) \quad (a) \quad |f| \leq K, \quad (b) \quad \left| \frac{\partial f}{\partial y} \right| \leq M \quad \text{for all } (x, y) \text{ in } R$$

then the initial value problem (1) has only one solution $y(x)$, which is defined at least for all x in that interval $|x - x_0| < \alpha$. This solution can then be obtained by Picard's iteration method, that is, the sequence $y_0, y_1, \dots, y_n, \dots$ where

$$y_n(x) = y_0 + \int_{x_0}^x f[t, y_{n-1}(t)] dt, \quad n = 1, 2,$$

converges to that solution $y(x)$.

Since the proofs of these theorems require familiarity with uniformly convergent series and other concepts to be considered later in this book, we shall not present these proofs at this time but refer the student to Ref. [B5] in Appendix 1. However, we want to include some remarks and examples which may be helpful for a good understanding of the two theorems.

Since $y' = f(x, y)$, the condition (2) implies that $|y'| \leq K$, that is, the slope of any solution curve $y(x)$ in R is at least $-K$ and at most K . Hence a solution curve which passes through the point (x_0, y_0) must lie in the shaded region in Fig. 56 bounded by the lines l_1 and l_2 whose slopes are $-K$ and K , respectively. Depending on the form of R , two different cases may arise. In the first case, shown in Fig. 56a, we have $b/K \geq a$ and therefore $\alpha = a$ in the existence theorem, which then asserts that the solution exists for all x between $x_0 - a$ and $x_0 + a$. In the second case, shown in Fig. 56b, we have $b/K < a$. Therefore $\alpha = b/K$, and all we can conclude from the theorems

¹⁵ A function $f(x, y)$ is said to be bounded when (x, y) varies in a region in the xy -plane, if there is a number K such that $|f| \leq K$ when (x, y) is in that region. For example, $f = x^2 + y^2$ is bounded, with $K = 2$ if $|x| < 1$ and $|y| < 1$. The function $f = \tan(x + y)$ is not bounded for $|x + y| < \pi/2$.

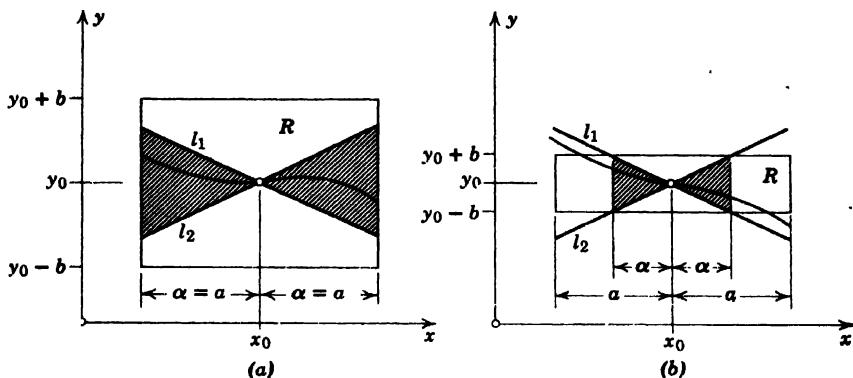


Fig. 56. The condition (2) of the existence theorem. (a) First case. (b) Second case.

is that the solution exists for all x between $x_0 - b/K$ and $x_0 + b/K$; for larger or smaller x 's the solution curve may leave the rectangle R , and since nothing is assumed about f outside R , nothing can be concluded about the solution for those corresponding values of x .

Example 1. Consider the problem

$$y' = 1 + y^2, \quad y(0) = 0 \quad (\text{cf. Ex. 1 in Sec. 1.11})$$

and take R : $|x| < 5$, $|y| < 3$. Then $a = 5$, $b = 3$, $|f| = |1 + y^2| \leq K = 10$, $|\partial f / \partial y| = 2|y| \leq M = 6$, $\alpha = b/K = 0.3 < a$. In fact, the solution $y = \tan x$ of the problem is discontinuous at $x = \pm\pi/2$, and there is no continuous solution valid in the entire interval $|x| < 5$ from which we started.

The conditions in the two theorems are sufficient conditions rather than necessary ones, and can be lessened. For example, by the mean value theorem of differential calculus we have

$$f(x, y_2) - f(x, y_1) = (y_2 - y_1) \frac{\partial f}{\partial y} \Big|_{y=\tilde{y}}$$

where (x, y_1) and (x, y_2) are assumed to be in R , and \tilde{y} is a suitable value between y_1 and y_2 . From this and (3b) it follows that

$$(4) \quad |f(x, y_2) - f(x, y_1)| \leq M |y_2 - y_1|,$$

and it can be shown that (3b) may be replaced by the weaker condition (4) which is known as a **Lipschitz condition**.¹⁶ However, continuity of $f(x, y)$ is not enough to guarantee the *uniqueness* of the solution. This may be illustrated by the following example.

¹⁶ RUDOLF LIPSCHITZ (1831–1903), German mathematician, who contributed also to algebra and differential geometry.

Example 2. The initial value problem

$$y' = \sqrt{|y|}, \quad y(0) = 0$$

has the two solutions

$$y \equiv 0 \quad \text{and} \quad y^* = \begin{cases} x^2/4 & \text{when } x \geq 0 \\ -x^2/4 & \text{when } x \leq 0 \end{cases}$$

although $f(x, y) = \sqrt{|y|}$ is continuous for all y . The Lipschitz condition (4) is violated in any region which includes the line $y = 0$, because for $y_1 = 0$ and positive y_2 we have

$$\frac{|f(x, y_2) - f(x, y_1)|}{|y_2 - y_1|} = \frac{\sqrt{y_2}}{y_2} = \frac{1}{\sqrt{y_2}}, \quad (\sqrt{y_2} > 0)$$

and this can be made as large as we please by choosing y_2 sufficiently small, while (4) requires that the quotient on the left does not exceed a fixed constant M .

1.13 Numerical Methods for Differential Equations of the First Order

We shall now consider some standard numerical methods for solving initial value problems of the form

$$(1) \quad y' = f(x, y), \quad y(x_0) = y_0.$$

Numerical methods for solving differential equations are methods which yield approximate numerical values of the solutions. They may be applied if a given differential equation cannot be solved by an exact method. This happens quite often in engineering mathematics, and these methods are therefore of great practical importance. They must also be applied if we want to solve a differential equation by the use of an electronic computer.

Many numerical methods for solving problems of the form (1) are such that we start from $y_0 = y(x_0)$ and proceed stepwise. That is, in the first step we compute a value y_1 of the solution of (1) at $x = x_0 + h$. In the second step we compute a value y_2 of that solution at $x = x_0 + 2h$, etc. Here h is some fixed number, e.g., 0.2 or 0.1 or 0.01, depending on the desired accuracy of the approximate values.

In each step the computation is done by the same formula. Such formulas can be obtained from the Taylor series

$$y(x + h) = y(x) + hy'(x) + \frac{h^2}{2} y''(x) + \dots$$

From (1) we have $y' = f$. Hence $y'' = f'$, etc., and the Taylor series becomes

$$(2) \quad y(x + h) = y(x) + hf + \frac{h^2}{2} f' + \frac{h^3}{6} f'' + \dots,$$

where the derivatives f', f'', \dots are evaluated at $[x, y(x)]$.

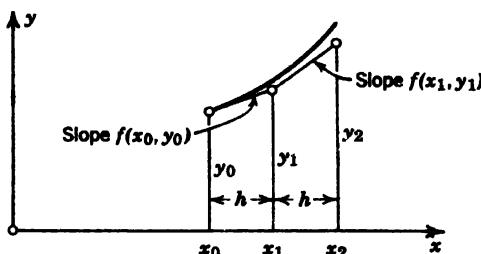


Fig. 57. Euler-Cauchy method.

Throughout this section we shall assume that f is such that (1) has a unique solution in some interval containing x_0 .

For small values of h the terms containing h^2, h^3, \dots in (2) will be very small; neglecting these terms, we obtain from (2) the approximation

$$y(x + h) \approx y(x) + hf.$$

Using this formula, we may now proceed stepwise as follows. In the first step we compute

$$y_1 = y_0 + hf(x_0, y_0),$$

which approximates $y(x_1) = y(x_0 + h)$. In the second step we compute

$$y_2 = y_1 + hf(x_1, y_1),$$

which approximates $y(x_2) = y(x_0 + 2h)$, etc., and in general

$$(3) \quad y_{n+1} = y_n + hf(x_n, y_n).$$

In this fashion we obtain approximate values to the solution $y(x)$ of the initial value problem (1) at $x = x_0, x_0 + h, x_0 + 2h, \dots$. This simple procedure is called the **Euler-Cauchy method**.¹⁷

Geometrically speaking, this method is an approximation of $y(x)$ by a polygon (Fig. 57).

The method can also be used in graphical form, that is, in (3) only $f(x_n, y_n)$ is computed while y_{n+1} is then obtained from y_n by means of a good graph.

The method is said to be a *first-order method* because in (2) we take only the term containing the first power of h . The omission of the further terms in (2) causes an error which is called the *truncation error* of the method (cf. Sec. 0.6). For small h , the third and higher powers of h will be small compared with h^2 contained in the first neglected term, and we therefore say that *the truncation error is of order h^2* . Note that another error of the method is due to the fact that, in (3), $f(x_n, y_n)$ is an approximation of $f[x_n, y(x_n)]$.

¹⁷ LEONHARD EULER (1707–1783) was an enormously creative Swiss mathematician. He contributed to almost all branches of mathematics and its application to physical problems. His important books on algebra and calculus include numerous results of his own research work. The great French mathematician, AUGUSTIN LOUIS CAUCHY (1789–1857), is the father of modern analysis. He exercised a great influence on the theory of infinite series, complex analysis, and differential equations.

Example 1. Apply the Euler-Cauchy method to the initial value problem

$$(4) \quad y' = x + y, \quad y(0) = 0,$$

choosing $h = 0.2$. Here $f(x, y) = x + y$, and (3) becomes

$$y_{n+1} = y_n + 0.2(x_n + y_n).$$

Table 3 shows a convenient arrangement of the numerical work.

Of course the exact solution of our problem is readily obtained from (4) in Sec. 1.7,

$$y = e^x - x - 1,$$

and we may determine the errors of our approximate values (cf. Table 8 in the current section).

Table 3. Euler-Cauchy Method Applied to (4).

n	x_n	y_n	$x_n + y_n$	$0.2(x_n + y_n)$	$y_n + 0.2(x_n + y_n)$
0	0	0	0	0	0
1	0.2	0	0.200	0.040	0.040
2	0.4	0.040	0.440	0.088	0.128
3	0.6	0.128	0.728	0.146	0.274
4	0.8	0.274	1.074	0.215	0.489
5	1.0	0.489			

In any practical case, when the exact solution is unknown, an indication of the accuracy of the values computed by the Euler-Cauchy method can be obtained by applying the method twice, using h and $h/2$, and then comparing corresponding values of the two approximations.

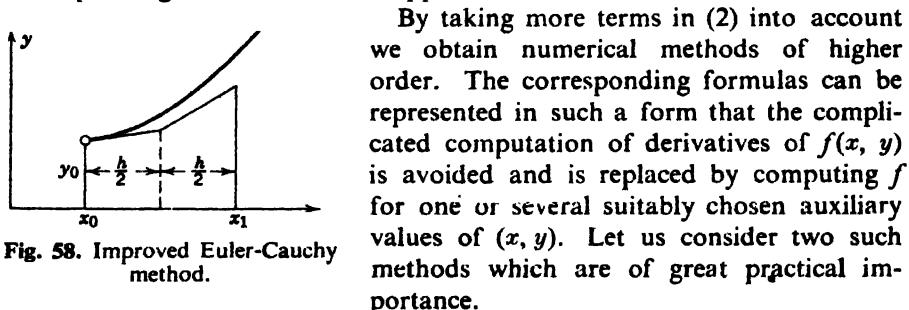


Fig. 58. Improved Euler-Cauchy method.

By taking more terms in (2) into account we obtain numerical methods of higher order. The corresponding formulas can be represented in such a form that the complicated computation of derivatives of $f(x, y)$ is avoided and is replaced by computing f for one or several suitably chosen auxiliary values of (x, y) . Let us consider two such methods which are of great practical importance.

The first method is the so-called **improved Euler-Cauchy method**. In each step of this method we first compute the auxiliary value

$$(5a) \quad y_{n+1}^* = y_n + hf(x_n, y_n),$$

and then the new value

$$(5b) \quad y_{n+1} = y_n + \frac{h}{2} [f(x_{n+1}, y_{n+1}^*) + f(x_n, y_n)].$$

Geometrically speaking, we may say that we approximate $y(x)$ for $x \geq x_n$ by the straight line through (x_n, y_n) with slope $f(x_n, y_n)$ until x reaches the value $x_n + \frac{1}{2}h$, and then we continue along the line with slope $f(x_{n+1}, y_{n+1}^*)$; cf. Fig. 58 (where $n = 0$),

It can be shown that the truncation error of the method is of the order h^3 (cf. Ref. [B7] in Appendix 1) and the improved Euler-Cauchy method is, therefore, a second-order method.

Example 2. Apply the improved Euler-Cauchy method to the initial value problem (4) in Ex. 1, choosing $h = 0.2$ as before. Here (5a) becomes

$$y_{n+1}^* = y_n + 0.2(x_n + y_n),$$

and (5b) becomes

$$y_{n+1} = y_n + 0.1[(x_{n+1} + y_{n+1}^*) + (x_n + y_n)].$$

Since in the present case the expression for y_{n+1}^* is very simple, it pays to insert it into the formula for y_{n+1} . Simplifying the resulting expression, we readily obtain

$$y_{n+1} = 0.12x_n + 0.1x_{n+1} + 1.22y_n.$$

The computations are shown in Table 4, and we see that the values are more accurate than those in Ex. 1 (cf. Table 8).

Table 4. Improved Euler-Cauchy Method Applied to (4).

n	x_n	y_n	$0.12x_n$	$0.1x_{n+1}$	$1.22y_n$	y_{n+1}
0	0	0	0	0.0200	0	0.0200
1	0.2	0.0200	0.0240	0.0400	0.0244	0.0884
2	0.4	0.0884	0.0480	0.0600	0.1078	0.2158
3	0.6	0.2158	0.0720	0.0800	0.2633	0.4153
4	0.8	0.4153	0.0960	0.1000	0.5067	0.7027
5	1.0	0.7027				

A still more accurate method is the so-called **Runge-Kutta method**,¹⁸ which is of great practical importance. In each step of this method we first compute four auxiliary quantities

$$(6a) \quad \begin{aligned} A_n &= hf(x_n, y_n), & B_n &= hf\left(x_n + \frac{h}{2}, y_n + \frac{A_n}{2}\right), \\ C_n &= hf\left(x_n + \frac{h}{2}, y_n + \frac{B_n}{2}\right), & D_n &= hf(x_{n+1}, y_n + C_n), \end{aligned}$$

and then the new value

$$(6b) \quad y_{n+1} = y_n + \frac{1}{6}(A_n + 2B_n + 2C_n + D_n).$$

It can be shown that the truncation error of the method is of the order h^5 (cf. Ref. [A15], Appendix 1) and the method is, therefore, a fourth-order method. Note that if f depends only on x , this method reduces to Simpson's rule of integration (Sec. 0.8).

The computations may be arranged as shown in Table 5.

¹⁸CARL RUNGE (1856–1927) and WILHELM KUTTA (1867–1944), German mathematicians.

Table 5. Runge-Kutta Method, First Step.

n	x	y		
0	x_0	y_0	$\frac{A_0}{2}$	A_0
	$x_0 + \frac{h}{2}$	$y_0 + \frac{A_0}{2}$	$\frac{B_0}{2}$	$2B_0$
	$x_0 + \frac{h}{2}$	$y_0 + \frac{B_0}{2}$	C_0	$2C_0$
	$x_0 + h$	$y_0 + C_0$		D_0
				Sum
				$K_0 = \frac{\text{Sum}}{6}$
		x_1	$y_1 (= y_0 + K_0)$	

Example 3. Apply the Runge-Kutta method to the initial value problem (4) in Ex. 1, choosing $h = 0.2$, as before. Here $f(x, y) = x + y$, and (6a) becomes

$$A_n = 0.2(x_n + y_n),$$

$$B_n = 0.2\left(x_n + 0.1 + y_n + \frac{A_n}{2}\right),$$

$$C_n = 0.2\left(x_n + 0.1 + y_n + \frac{B_n}{2}\right),$$

$$D_n = 0.2(x_n + 0.2 + y_n + C_n).$$

Table 6 shows the computations in the first step (arranged as in Table 5).

Table 6. Example 3, First Step.

n	x	y		
0	0	0	0	0
	0.1	0	0.0100	0.0400
	0.1	0.0100	0.0220	0.0440
	0.2	0.0220		0.0444
				0.1284
				$K_0 = 0.0214$
		0.2	0.0214	

We note that in the present case, where A_n , B_n , C_n , and D_n are very simple expressions it is worthwhile to insert A_n into B_n , finding

$$B_n = 0.22(x_n + y_n) + 0.02,$$

then insert this expression into C_n , finding

$$C_n = 0.222(x_n + y_n) + 0.022,$$

and finally insert this expression into D_n , finding

$$D_n = 0.2444(x_n + y_n) + 0.0444.$$

If we use these expressions, (6b) becomes

$$(7) \quad y_{n+1} = y_n + 0.2214(x_n + y_n) + 0.0214.$$

Table 7 shows the corresponding computations, and we see that the values are much more accurate than those in Ex. 1 and 2.

Table 7. Runge-Kutta Method Applied to (4); Computations by the Use of (7).

n	x_n	y_n	$x_n + y_n$	$0.2214(x_n + y_n)$	$0.2214(x_n + y_n) + 0.0214$
0	0	0	0	0	0.021 400
1	0.2	0.021 400	0.221 400	0.049 018	0.070 418
2	0.4	0.091 818	0.491 818	0.108 889	0.130 289
3	0.6	0.222 107	0.822 107	0.182 014	0.203 414
4	0.8	0.425 521	1.225 521	0.271 330	0.292 730
5	1.0	0.718 251			

Table 8. Comparison of the Accuracy of the Three Methods under Consideration in the Case of the Initial Value Problem (4), with $h = 0.2$.

x	$y = e^x - x - 1$	Absolute value of error		
		Euler-Cauchy (Table 3)	Improved Euler- Cauchy (Table 4)	Runge-Kutta (Table 7)
0.2	0.021 403	0.021	0.0014	0.000 003
0.4	0.091 825	0.052	0.0034	0.000 007
0.6	0.222 119	0.094	0.0063	0.000 011
0.8	0.425 541	0.152	0.0102	0.000 020
1.0	0.718 282	0.229	0.0156	0.000 031

Problems

1. Repeat the computations in Ex. 1, choosing $h = 0.1$ (instead of $h = 0.2$), and note that the errors of the values thus obtained are about 50% of those in Ex. 1 (which are listed in Table 8).
2. Same task as in Prob. 1, but with $h = 0.01$ (twenty steps). Compare the error of the value corresponding to $x = 0.2$ with that in Ex. 1.
3. Repeat the computations in Ex. 2, choosing $h = 0.1$, and note that the errors of the values thus obtained are about 25% of those in Ex. 2.
4. Same task as in Prob. 3, but with $h = 0.05$ (eight steps). Compare the error of the value corresponding to $x = 0.4$ with those in Ex. 2 and Prob. 3.

Apply the Euler-Cauchy method to the following differential equations. (Carry out ten steps.)

$$5. y' = y, y(0) = 1, h = 0.1$$

$$7. y' = .2xy + 1, y(0) = 0, h = 0.1$$

$$6. y' = y, y(0) = 1, h = 0.01$$

$$8. y' = 2xy, y(0) = 1, h = 0.1$$

9. Compute $y = e^x$ for $x = 0, 0.1, 0.2, \dots, 1.0$ by applying the Runge-Kutta method to the initial value problem in Prob. 5. Show that the first five decimal places of the results are correct (cf. Table 1 in Sec. 0.1).
10. In Prob. 9 insert A_n into B_n , then B_n into C_n , etc., and show that the resulting formula becomes

$$y_{n+1} = 1.105\,170\,833y_n.$$

Repeat the computations in Prob. 9, using this formula.

11. Show that for $x = 0.1$

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} = 1.105\,170\,833\dots,$$

the constant in Prob. 10. What is the reason for this result?

12. Using the idea of Prob. 11, show that a more accurate recursion formula for computing $y = e^x, x = 0, 0.1, 0.2, \dots$ is

$$y_{n+1} = 1.105\,170\,917y_n.$$

Compute nine decimal places of e^x for $x = 0.1, 0.2, 0.3$, using this formula, and compare the result with the accurate values 1.105 170 918, 1.221 402 758, 1.349 858 808.

13. Apply the Euler-Cauchy method to the initial value problem $y' = 1 + y^2, y(0) = 0$, choosing $h = 0.1$. Compute five steps and compare the results with the exact values (cf. Table 1 in Sec. 0.1).
14. Apply the improved Euler-Cauchy method to the initial value problem in Prob. 13, choosing $h = 0.1$. Compute two steps and compare the errors with those of the corresponding values in Prob. 13.
15. Apply the Runge-Kutta method to the initial value problem in Prob. 13, choosing $h = 0.1$. Compute two steps. Compare the results with the exact values 0.100 334 672 and 0.202 710 036.
16. Apply the improved Euler-Cauchy method to the initial value problem $y' = 2x, y(0) = 0$, choosing $h = 0.1$. Why are the errors zero?

Chapter 2

Ordinary Linear Differential Equations

The ordinary differential equations may be divided into two large classes, the so-called linear equations and the nonlinear equations. While nonlinear equations (of the second and higher orders) are rather difficult, linear equations are much simpler in many respects, because various properties of their solutions can be characterized in a general way, and standard methods are available for solving many of these equations. In the present chapter we shall consider linear differential equations and their applications. These equations play an important role in engineering mathematics, for example, in connection with mechanical vibrations and electric circuits and networks.

Much space will be devoted to linear equations of the second order. This seems to be justified by the following facts. First of all, second-order equations are the most important ones from the practical point of view. Furthermore, the necessary theoretical considerations will become easier if we first concentrate upon second-order equations. Once the student understands what is needed to handle second-order equations, he will easily become familiar with the generalizations of the concepts, methods, and results to differential equations of the third and higher orders.

Important engineering applications in connection with mechanical and electrical oscillations will be considered in Secs. 2.6 and 2.13–2.15.

Numerical methods for solving second-order differential equations are included in Sec. 2.17.

(Legendre's, Bessel's, and the hypergeometric equations will be considered in Chap. 3.)

Prerequisite for this chapter: Chap. 1, in particular Secs. 1.7–1.9.

Sections which may be omitted in a shorter course: 2.8–2.10, 2.15–2.17.

References: Appendix 1, Part B.

Answers to problems: Appendix 2.

2.1 Homogeneous Linear Equations of the Second Order

The student has already met with linear differential equations of the first order (Secs. 1.7–1.9), and we shall now define and consider linear equations of the second order.

A second-order differential equation is said to be **linear** if it can be written in the form

$$(1) \quad y'' + f(x)y' + g(x)y = r(x).$$

The characteristic feature of this equation is that it is linear in y and its derivatives, while f , g , and r may be any given functions of x . Any equation of the second order which cannot be written in the form (1) is said to be **nonlinear**. For example

$$(2) \quad y'' + 4y = e^{-x} \sin x$$

and

$$(3) \quad x^2y'' + xy' + (x^2 - 1)y = 0$$

are linear equations, while

$$y''y + y' = 0$$

and

$$y'' + \sqrt{y} = 0$$

are nonlinear equations.

Linear differential equations of the second order play a basic role in many different engineering problems. We shall see that some of these equations are very simple because their solutions are elementary functions. Others are more complicated, their solutions being important higher functions, such as Bessel and hypergeometric functions.

Our further consideration will show that equation (1) has somewhat simpler properties if the function $r(x)$ on the right-hand side is identically zero. Then equation (1) becomes

$$(4) \quad y'' + f(x)y' + g(x)y = 0$$

and is said to be **homogeneous**. If $r(x) \not\equiv 0$, equation (1) is said to be **nonhomogeneous**. These notions are quite similar to those in the case of linear equations of the first order in Sec. 1.7.

For example, (2) is nonhomogeneous and (3) is homogeneous.

The functions f and g in (1) and (4) are called the **coefficients** of the equations.

In all our considerations x will be assumed to vary in some arbitrary fixed range, for example, in some finite interval or on the entire x -axis. All the assumptions and statements will refer to such a fixed range, which need not be specified in each case.

A function

$$y = \phi(x)$$

is called a **solution** of a (linear or nonlinear) differential equation of the second order *on some interval* (perhaps infinite), if $\phi(x)$ is defined and twice differentiable throughout that interval and is such that the equation becomes an identity when we replace the unspecified function y and its derivatives in the equation by ϕ and its derivatives.

This definition is analogous to that in the case of first-order equations in Sec. 1.1.

Example 1. The functions $y = \cos x$ and $y = \sin x$ are solutions of the homogeneous linear differential equation

$$y'' + y = 0$$

for all x , since

$$(\cos x)'' + \cos x = -\cos x + \cos x = 0$$

and similarly for $y = \sin x$. We can even go an important step further. If we multiply the first solution by a constant, for example, by 3, the resulting function $y = 3 \cos x$ is also a solution, because

$$(3 \cos x)'' + 3 \cos x = 3[(\cos x)'' + \cos x] = 0.$$

It is clear that instead of 3 we may choose any other constant, for example, -5 or $\frac{2}{3}$. We may even multiply $\cos x$ and $\sin x$ by different constants, say, by 2 and -8 , respectively, and add the resulting functions, having

$$y = 2 \cos x - 8 \sin x,$$

and this function is also a solution of our homogeneous equation for all x , because

$$(2 \cos x - 8 \sin x)'' + 2 \cos x - 8 \sin x = 2[(\cos x)'' + \cos x] - 8[(\sin x)'' + \sin x] = 0.$$

This example illustrates the very important fact that we may obtain new solutions of the homogeneous linear equation (4) from unknown solutions by multiplication by constants and by addition. Of course, this is a great practical and theoretical advantage, because this property will enable us to generate more complicated solutions from simple solutions. We may characterize this basic property as follows.

Fundamental theorem 1. *If a solution of the homogeneous linear differential equation (4) is multiplied by any constant, the resulting function is also a solution of (4). If two solutions of that equation are added, the resulting sum is also a solution of that equation.¹*

Proof. We assume that $\phi(x)$ is a solution of (4) and show that then $y = c\phi(x)$ is also a solution of (4). If we substitute $y = c\phi(x)$ into (4), the left-hand side of (4) becomes

$$(c\phi)'' + f(c\phi)' + gc\phi = c[\phi'' + f\phi' +$$

¹ This theorem is sometimes called the **superposition principle** or **linearity principle**.

Since ϕ satisfies (4), the expression in brackets is zero, and the first statement of the theorem is proved. The proof of the last statement is quite simple, too, and is left to the reader.

The student should know this highly important theorem as well as he knows his multiplication table, but he should not forget that *the theorem does not hold for nonhomogeneous linear equations or nonlinear equations*, as may be illustrated by the following two examples.

Example 2. Substitution shows that the functions

$$y = 1 + \cos x \quad \text{and} \quad y = 1 + \sin x$$

are solutions of the nonhomogeneous linear equation

$$y'' + y = 1,$$

but the functions

$$2(1 + \cos x) \quad \text{and} \quad (1 + \cos x) + (1 + \sin x)$$

are not solutions of that equation.

Example 3. Substitution shows that the functions

$$y = x^2 \quad \text{and} \quad y = 1$$

are solutions of the nonlinear differential equation

$$y''y - xy' = 0,$$

but the functions

$$-x^2 \quad \text{and} \quad x^2 + 1$$

are not solutions of that equation.

Problems

Prove that the linear equations (1) (with $r(x) \neq 0$) and (4) have the following important properties.

1. $y \equiv 0$ is a solution of (4), but not of (1).
2. The sum $y = y_1 + y_2$ of two solutions y_1 and y_2 of (1) is *not* a solution of (1) but satisfies $y'' + fy' + gy = 2r$.
3. The difference $y = y_1 - y_2$ of two solutions y_1 and y_2 of (1) is a solution of (4).
4. If y_1 is a solution of (1), then $y = cy_1$ ($c \neq 1$) is *not* a solution of (1) but satisfies $y'' + fy' + gy = cr$.
5. The sum $y = y_1 + y_2$ of a solution y_1 of (1) and a solution y_2 of (4) is a solution of (1).
6. If y_1 and y_2 are, respectively, solutions of

$$y'' + fy' + gy = r_1 \quad \text{and} \quad y'' + fy' + gy = r_2,$$

then their sum $y = y_1 + y_2$ is a solution of

$$y'' + fy' + gy = r_1 + r_2.$$

Second-order differential equations reducible to the first order. The following problems illustrate that certain second-order differential equations can be reduced to the first order.

7. If in a second-order equation the dependent variable y does not appear explicitly, the equation is of the form $F(x, y', y'') = 0$. Show that by setting $y' = z$ this equation becomes a first-order equation in z , and from a solution z a solution y of the original equation can be obtained by integration.

Reduce to the first order and solve

8. $y'' = 2y'$

9. $y'' + 4y' = 0$

10. $xy'' = y'$

11. $xy'' + 2y' = 0$

12. $y'' = -2y' \tanh x$

13. $y'' = 1 + y'^2$

14. If in a second-order differential equation the independent variable x does not appear explicitly, the equation is of the form $F(y, y', y'') = 0$. Show that this equation can be reduced to a first-order equation in $z = y'$ in which y is the independent variable. Hint: using the chain rule of differentiation, show that $y'' = (dz/dy)x$.

Reduce to the first order and solve

15. $yy'' = 2y'^2$

16. $yy'' = y'^2$

17. $y'' + y'^2 = 0$

18. Show that the method in Prob. 7 yields the same answer to Prob. 17 as the method in Prob. 14.

19. Given $y'' + ay'^2 + f(y) = 0$ where a is a constant. Show that the substitutions $y' = z$, $v = z^2$ lead to the linear first-order equation

$$\frac{dv}{dy} + 2av + 2f(y) = 0.$$

20. A particle moves on a straight line so that its acceleration equals its velocity. At the time $t = 0$ its displacement from the origin is 1 ft and its velocity is 3 ft/sec. Find its position and velocity when $t = 5$ sec.
21. Same task as in Prob. 20, assuming that the acceleration equals minus the square of the velocity, the other data being as before.
22. A particle moves on a straight line so that its acceleration is equal to the reciprocal of its velocity. At the time $t = 0$ the displacement from the origin is 4 ft and the velocity is 3 ft/sec. Find its displacement and velocity when $t = 2$ sec.
23. Same task as in Prob. 22, assuming that the displacements at $t = 0$ and $t = 8$ are 0 and $64/3$ ft, respectively, the other data being as before.
24. Find the curve $y(x)$ through the origin for which $y'' = y'$ and the angle between the tangent at the origin and the x -axis equals 45° .
25. Find the curve $y(x)$ through the points $(0, 0)$ and $(1, 1)$ for which $2y''y = y'^2$.

2.2 Homogeneous Second-Order Equations with Constant Coefficients

We shall now consider linear differential equations of the second order whose coefficients are constant. These equations have important engineering applications, especially in connection with mechanical and electrical vibrations, as will be demonstrated in Secs. 2.6, 2.13, and 2.14.

We start with homogeneous equations, that is, equations of the form

(1) $y'' + ay' + by = 0$

where a and b are constants.

We shall assume that a and b are real and that the range of x considered is the entire x -axis.

How to solve equation (1)? We remember that the solution of the *first-order* linear homogeneous equation with constant coefficients

$$y' + ky = 0$$

is an exponential function, namely,

$$y = ce^{-kx}.$$

It seems natural to suppose that

$$(2) \quad y = e^{\lambda x}$$

may be a solution of (1) if λ is properly chosen. Substituting (2) and its derivatives

$$y' = \lambda e^{\lambda x} \quad \text{and} \quad y'' = \lambda^2 e^{\lambda x}$$

into our equation (1), we obtain

$$(\lambda^2 + a\lambda + b)e^{\lambda x} = 0.$$

So (2) is a solution of (1), if λ is a solution of the quadratic equation

$$(3) \quad \lambda^2 + a\lambda + b = 0.$$

This equation is called the **characteristic equation** (or *auxiliary equation*) of (1). Its roots are

$$(4) \quad \lambda_1 = \frac{1}{2}(-a + \sqrt{a^2 - 4b}), \quad \lambda_2 = \frac{1}{2}(-a - \sqrt{a^2 - 4b}).$$

From our derivation it follows that the functions

$$(5) \quad y_1 = e^{\lambda_1 x} \quad \text{and} \quad y_2 = e^{\lambda_2 x}$$

are solutions of (1). The student may check this result by substituting (5) into (1).

From elementary algebra we know that, since a and b are real, the characteristic equation may have

(Case I) two distinct real roots,

(Case II) two complex conjugate roots, or

(Case III) a real double root.

These cases will be discussed separately. For the time being let us illustrate each case by a simple example.

Example 1. Find solutions of the equation

$$y'' + y' - 2y = 0.$$

The characteristic equation is

$$\lambda^2 + \lambda - 2 = 0.$$

The roots are 1 and -2, and we obtain the two solutions

$$y_1 = e^x \quad \text{and} \quad y_2 = e^{-2x}.$$

Example 2. Find solutions of the equation

$$y'' + y = 0.$$

The characteristic equation is

$$\lambda^2 + 1 = 0.$$

The roots are $i (= \sqrt{-1})$ and $-i$, and we obtain the two solutions

$$y_1 = e^{ix} \quad \text{and} \quad y_2 = e^{-ix}.$$

In Sec. 2.4 we shall see that it is quite simple to obtain real solutions from such complex solutions.

Example 3. Find solutions of the equation

$$y'' - 2y' + y = 0.$$

The characteristic equation

$$\lambda^2 - 2\lambda + 1 = 0$$

has the real double root 1, and we obtain at first only one solution

$$y_1 = e^x.$$

The case of a real double root will be discussed in Sec. 2.5.

Problems

Find solutions of the following differential equations.

- | | |
|-------------------------|---------------------------|
| 1. $y'' - 9y = 0$ | 2. $y'' - y = 0$ |
| 3. $y'' + y' - 2y = 0$ | 4. $y'' + 2y' = 0$ |
| 5. $y'' - y' = 0$ | 6. $y'' - 3y' + 2y = 0$ |
| 7. $y'' + y = 0$ | 8. $y'' + \omega^2 y = 0$ |
| 9. $y'' + 4y' + 5y = 0$ | 10. $y'' + 2y' + 5y = 0$ |

11. Show that $y = \cosh 3x$ and $y = \sinh 3x$ are solutions of $y'' - 9y = 0$. How can these solutions be obtained from those in the answer to Prob. 1?
12. Show that (5) are solutions of (1) if and only if $a = -\lambda_1 - \lambda_2$ and $b = \lambda_1 \lambda_2$.

Find a differential equation of the form (1) for which the following functions are solutions.

- | | | | |
|-------------------------|-------------------------|----------------------------------|-----------------|
| 13. e^{-x}, e^{-2x} | 14. e^{2x}, e^{-2x} | 15. e^{3x}, e^{4x} | 16. $1, e^{2x}$ |
| 17. e^{3ix}, e^{-3ix} | 18. e^{2ix}, e^{-2ix} | 19. $e^{(-3+4i)x}, e^{(-3-4i)x}$ | |

20. Show that the cases I, II, and III occur if and only if $a^2 > 4b$, $a^2 = 4b$, and $a^2 < 4b$, respectively.

2.3 General Solution. Fundamental System

In our further discussion of linear equations with constant coefficients, it will be necessary to know about some important general properties of the solutions of linear differential equations to be considered in this section. We first introduce the following two concepts. .

A solution of a differential equation of the second order (linear or not) is called a general solution if it contains two arbitrary² independent constants. Here independence means that the same solution cannot be reduced to a form containing only one arbitrary constant or none. If we assign definite values to those two constants, then the solution so obtained is called a particular solution.

These notions are obvious generalizations of those in the case of first-order differential equations (Sec. 1.1).

It will be shown later (in Secs. 2.8 and 2.11) that if a differential equation of the second order is linear (homogeneous or not), then a general solution includes all solutions of the equation; any solution containing no arbitrary constants can be obtained by assigning definite values to the arbitrary constants in that general solution (cf. also the general remark in Sec. 1.1, p. 43).

In the present section we shall consider the homogeneous linear equation

$$(1) \quad y'' + f(x)y' + g(x)y = 0,$$

and we want to show that a general solution of such an equation can be readily obtained if two suitable solutions y_1 and y_2 are known.

Clearly, if $y_1(x)$ and $y_2(x)$ are solutions of (1) on some interval I , then, by Fundamental Theorem 1 in Sec. 2.1,

$$(2) \quad y(x) = c_1y_1(x) + c_2y_2(x)$$

where c_1 and c_2 are arbitrary constants, will be a solution of (1) on I . Since it contains two arbitrary constants, it will be a general solution of (1), provided it cannot be reduced to an expression containing less than two arbitrary constants.

Under what conditions will such a reduction be possible? To consider this question, we first introduce the following basic concepts, which are important in other considerations, too.

Two functions $y_1(x)$ and $y_2(x)$ are said to be linearly dependent on an interval I where both functions are defined, if they are proportional on I , that is, if³

$$(3) \quad (a) \quad y_1 = ky_2 \quad \text{or} \quad (b) \quad y_2 = ly_1$$

holds for all x on I ; here k and l are numbers, zero or not. If the functions are not proportional on I , they are said to be linearly independent on I .

Note that the concepts of linear dependence and independence always refer to an interval, but not merely to a single point.

~~If at least one of the functions y_1 and y_2 is identically zero on I , then the functions are linearly dependent on I . In any other case the functions are~~

² The range of the constants may have to be restricted in some cases to avoid imaginary expressions or other degeneracies.

³ Note that if (3a) holds for a $k \neq 0$, we may divide by k , finding $y_2 = y_1/k$, so that in this case (3a) implies (3b) (with $l = 1/k$). However, if $y_1 \equiv 0$, then (3a) holds for $k = 0$, but does not imply (3b); in fact, in this case (3b) will not hold unless $y_2 \equiv 0$, too.

linearly dependent on I if, and only if, the quotient y_1/y_2 is constant on I. Hence, if y_1/y_2 depends on x on I, then y_1 and y_2 are linearly independent on I.

In fact, if $y_1 \equiv 0$, then (3a) holds with $k = 0$; and if $y_2 \equiv 0$, then (3b) holds with $l = 0$. This proves the first statement, and the other statements are immediate consequences of (3), too. Of course, in the last statement, we admit that y_1/y_2 may become infinite, namely, for those x for which $y_2 = 0$.

Example 1. The functions

$$y_1 = 3x \quad \text{and} \quad y_2 = x$$

are linearly dependent on any interval, because $y_1/y_2 = 3 = \text{const}$. The functions

$$y_1 = x^2 \quad \text{and} \quad y_2 = x$$

are linearly independent on any interval because $y_1/y_2 = x \neq \text{const}$, and so are the functions

$$y_1 = x \quad \text{and} \quad y_2 = x + 1.$$

If the solutions y_1 and y_2 in (2) are linearly dependent on I, then (3a) or (3b) holds and we see that we may then reduce (2) to one of the forms

$$y = Ay_2 \quad (A = c_1 k + c_2)$$

and

$$y = By_1 \quad (B = c_1 + lc_2)$$

containing only one arbitrary constant. Hence in this case (2) is certainly not a general solution of (1) on I.

On the other hand, if y_1 and y_2 are linearly independent on I, then they are not proportional, and no such reduction can be performed. Hence, in this case, (2) is a general solution of (1) on I.

Two linearly independent solutions of (1) on I are called a **fundamental system** of solutions on I. Using this notion, we may formulate our result as follows.

Theorem 1. *The solution*

$$y(x) = c_1 y_1(x) + c_2 y_2(x) \quad (c_1, c_2 \text{ arbitrary})$$

is a general solution of the differential equation (1) on an interval I of the x-axis if, and only if, the functions y_1 and y_2 constitute a fundamental system of solutions of (1) on I.

y_1 and y_2 constitute such a fundamental system if, and only if, their quotient⁴ y_1/y_2 is not constant on I but depends on x .

Example 2. The functions

$$y_1 = e^x \quad \text{and} \quad y_2 = e^{-2x}$$

considered in Ex. 1, Sec. 2.2, are solutions of

$$y'' + y' - 2y = 0.$$

⁴ Of course, here we exclude the trivial case $y_2 \equiv 0$ in which y_1 and y_2 do not constitute a fundamental system, and we also admit that y_1/y_2 may become infinite (namely, for those values of x for which $y_2 = 0$).

Since y_1/y_2 is not constant, these solutions constitute a fundamental system, and the corresponding general solution for all x is

$$y = c_1 y_1 + c_2 y_2 = c_1 e^x + c_2 e^{-2x}.$$

Example 3. The functions

$$y_1 = e^x \quad \text{and} \quad y_2 = 3e^x$$

are solutions of the equation in Ex. 2, but since their quotient is constant, they do not form a fundamental system.

We are now prepared for continuing our discussion of differential equations with constant coefficients, and shall, therefore, postpone further theoretical considerations until Sec. 2.8.

Problems

Are the following functions linearly dependent or independent on the given interval?

- | | |
|---|---|
| 1. e^x, e^{-x} ($-\infty < x < \infty$) | 2. $e^{\lambda_1 x}, e^{\lambda_2 x}, \lambda_1 \neq \lambda_2$ ($0 < x < 1$) |
| 3. x, x^2 ($-1 < x < 1$) | 4. $\cos x, \sin x$ ($-\pi < x < \pi$) |
| 5. $x + 1, 2x + 2$ ($0 < x < 1$) | 6. $e^{\lambda x}, xe^{\lambda x}$ ($-\infty < x < \infty$) |
| 7. $ x x, x^2$ ($0 < x < 1$) | 8. $ x x, x^2$ ($-1 < x < 1$) |
| 9. $\ln x, (\ln x)^2$ ($0 < x < 1$) | 10. $\ln x, \ln x^2$ ($x > 0$) |

11. If $y_1(x)$ and $y_2(x)$ are linearly dependent on an interval I , show that these functions are also linearly dependent on any subinterval of I .
12. Show by an example that two functions $y_1(x)$ and $y_2(x)$, which are linearly independent on an interval I , may be linearly dependent on a subinterval of I .

Find a general solution of the following differential equations.

- | | |
|--------------------------|--------------------------|
| 13. $y'' - y = 0$ | 14. $y'' - 3y' + 2y = 0$ |
| 15. $y'' - y' - 12y = 0$ | 16. $y'' + y' - 6y = 0$ |
| 17. $y'' + 4y = 0$ | 18. $y'' + 25y = 0$ |
| 19. $y'' - 2y' + 2y = 0$ | 20. $y'' + 4y' + 5y = 0$ |

Find a differential equation of the form $y'' + ay' + by = 0$ for which the given functions constitute a fundamental system of solutions. Test linear independence by means of Theorem 1.

- | | | |
|-------------------------------------|-----------------------|--------------------------------|
| 21. e^{2x}, e^{2x} | 22. e^{-2x}, e^{2x} | 23. $e^{(-3+i)x}, e^{(-3-i)x}$ |
| 24. $e^{i\omega x}, e^{-i\omega x}$ | 25. $1, e^{-4x}$ | 26. $e^{(2+i)x}, e^{(2-i)x}$ |

2.4 Complex Roots ✓ of the Characteristic Equation. Initial Value Problem

In Sec. 2.2 we saw that solutions of the homogeneous linear equation with constant coefficients

$$(1) \quad y'' + ay' + by = 0 \quad (a, b \text{ real})$$

can be obtained by a purely algebraic procedure, namely, by determining the roots λ_1, λ_2 of the corresponding characteristic equation

$$(2) \quad \lambda^2 + a\lambda + b = 0,$$

and these solutions are of the form

$$(3) \quad y_1 = e^{\lambda_1 x}, \quad y_2 = e^{\lambda_2 x}.$$

If $\lambda_1 \neq \lambda_2$, the quotient y_1/y_2 is not constant, and from Theorem 1 in Sec. 2.3 it follows that these solutions constitute a fundamental system for all x . The corresponding general solution is

$$(4) \quad y = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}.$$

If these distinct roots are real (Case I in Sec. 2.2), then the solutions (3) are real. However, if λ_1 and λ_2 are complex conjugate (Case II in Sec. 2.2), say,

$$\lambda_1 = p + iq, \quad \lambda_2 = p - iq \quad (p, q \text{ real}, q \neq 0)$$

then the solutions (3) are complex,

$$y_1 = e^{(p+iq)x}, \quad y_2 = e^{(p-iq)x},$$

and it is of great practical interest to obtain real solutions from these complex solutions.

This can be done by applying the *Euler formulas*⁵

$$e^{i\theta} = \cos \theta + i \sin \theta, \quad e^{-i\theta} = \cos \theta - i \sin \theta,$$

taking $\theta = qx$. In this way y_1 becomes

$$y_1 = e^{(p+iq)x} = e^{px} e^{iqx} = e^{px} (\cos qx + i \sin qx),$$

and by applying the second Euler formula,

$$y_2 = e^{(p-iq)x} = e^{px} e^{-iqx} = e^{px} (\cos qx - i \sin qx).$$

From these two expressions for y_1 and y_2 we now obtain

$$\frac{1}{2}(y_1 + y_2) = e^{px} \cos qx,$$

$$\frac{1}{2i}(y_1 - y_2) = e^{px} \sin qx.$$

We see that these two functions on the right-hand sides are real, and from Theorem 1 in Sec. 2.1 it follows that they are solutions of the differential equation (1). Furthermore, since their quotient is not a constant, they are linearly independent on any interval and thus constitute a fundamental system of solutions of (1) on the entire x -axis. The corresponding general solution is

$$(5) \quad y(x) = e^{px}(A \cos qx + B \sin qx)$$

where A and B are arbitrary constants.

⁵ Students not familiar with these important formulas from elementary calculus will find a proof in Sec. 10.5. Of course, once we obtained (5) we may verify directly that it satisfies (1) in the present case.

Example 1. Find a general solution of the equation

$$y'' - 2y' + 10y = 0.$$

The characteristic equation

$$\lambda^2 - 2\lambda + 10 = 0$$

has the roots $\lambda_1 = p + iq = 1 + 3i$ and $\lambda_2 = p - iq = 1 - 3i$. Thus $p = 1$, $q = 3$, and the answer is

$$y = e^x(A \cos 3x + B \sin 3x).$$

Example 2. A general solution of the equation

$$y'' + \omega^2 y = 0$$

(ω constant, not zero)

is

$$y = A \cos \omega x + B \sin \omega x.$$

In many practical cases it is required to find a particular solution $y(x)$ of (1) which has a given value K at a point $x = x_0$ and whose derivative has a given value L at $x = x_0$; that is,

$$y = K \quad \text{and} \quad y' = L \quad \text{when } x = x_0,$$

or more briefly,

$$(6) \quad y(x_0) = K, \quad y'(x_0) = L.$$

The equation (1), together with the two conditions (6), constitutes what is called an **initial value problem**. To solve such a problem we must find a particular solution of (1) satisfying (6). We shall see later (in Sec. 2.8) that such a problem has a unique solution; practically speaking, this means that by the use of the conditions (6) we can assign definite values to the arbitrary constants in a general solution of (1), the result being a unique particular solution satisfying (1) and (6). Let us illustrate the simple practical procedure, which is quite similar to that in the case of a differential equation of the first order.

Example 3. Solve the initial value problem

$$y'' - 2y' + 10y = 0, \quad y(0) = 4, \quad y'(0) = 1.$$

A fundamental system of solutions is

$$e^x \cos 3x \quad \text{and} \quad e^x \sin 3x$$

(cf. Ex. 1). The corresponding general solution is

$$y(x) = e^x(A \cos 3x + B \sin 3x),$$

and $y(0) = A$. By differentiation,

$$y'(x) = e^x[(A + 3B) \cos 3x + (B - 3A) \sin 3x],$$

and $y'(0) = A + 3B$. From this and the given initial conditions,

$$y(0) = A = 4,$$

$$y'(0) = A + 3B = 1.$$

Hence $A = 4$, $B = -1$, and the answer is

$$(7) \quad y = e^x(4 \cos 3x - \sin 3x).$$

Example 4. Solve the initial value problem

$$y'' + y' - 2y = 0, \quad y(0) = 3, \quad y'(0) = 0.$$

A general solution is

$$y(x) = c_1 e^x + c_2 e^{-2x},$$

and

$$y'(x) = c_1 e^x - 2c_2 e^{-2x}.$$

From this and the initial conditions,

$$y(0) = c_1 + c_2 = 3, \quad y'(0) = c_1 - 2c_2 = 0.$$

Hence $c_1 = 2$, $c_2 = 1$, and the answer is

$$y = 2e^x + e^{-2x}$$

Problems

Verify that the given functions are solutions of the corresponding equations.

1. $y = A \cos \omega x + B \sin \omega x, \quad y'' + \omega^2 y = 0$
2. $y = -3 \cos 4x + 0.5 \sin 4x, \quad y'' + 16y = 0$
3. $y = e^{-x}(A \cos 3x + B \sin 3x), \quad y'' + 2y' + 10y = 0$
4. $y = e^{-\alpha x}(2 \cos x - \sin x), \quad y'' + 2\alpha y' + (\alpha^2 + 1)y = 0$
5. $y = e^{-x/2}(\cos \sqrt{2}x + 3 \sin \sqrt{2}x), \quad y'' + y' + 2.25y = 0$

Verify that the following functions are solutions of the corresponding equations and represent them in the form (5).

6. $y = c_1 e^{2ix} + c_2 e^{-2ix}, \quad y'' + 4y = 0$
7. $y = c_1 e^{i\pi x} + c_2 e^{-i\pi x}, \quad y'' + \pi^2 y = 0$
8. $y = c_1 e^{(-1+i)x} + c_2 e^{(-1-i)x}, \quad y'' + 2y' + 2y = 0$
9. $y = c_1 e^{(3+4i)x} + c_2 e^{(3-4i)x}, \quad y'' - 6y' + 25y = 0$
10. $y = c_1 e^{(-\alpha+i\omega)x} + c_2 e^{(-\alpha-i\omega)x}, \quad y'' + 2\alpha y' + (\alpha^2 + \omega^2)y = 0$

11. Assuming that (5) is a solution of (1), express a and b in terms of p and q .

Using the result of Prob. 11, derive equations of the form (1) having the following general solutions:

- | | |
|------------------------------------|--|
| 12. $y = A \cos 2x + B \sin 2x$ | 13. $y = e^{-x}(A \cos x + B \sin x)$ |
| 14. $y = e^x(A \cos x + B \sin x)$ | 15. $y = e^{-3x}(A \cos \sqrt{3}x + B \sin \sqrt{3}x)$ |

Find a general solution of the following equations.

- | | |
|--------------------------|--|
| 16. $y'' - 2y' + 5y = 0$ | 17. $y'' + 0.25y = 0$ |
| 18. $y'' + 2y' + 5y = 0$ | 19. $y'' + 0.2y' + 0.26y = 0$ |
| 20. $y'' - 0.25y = 0$ | 21. $y'' + 6y' + 8y = 0$ |
| 22. $y'' + 2y' + 4y = 0$ | 23. $y'' + 2\alpha y' + (\alpha^2 + 1)y = 0$ |
| 24. $y'' - 2y' = 0$ | 25. $y'' + 5y = 0$ |
| 26. $y'' - 5y = 0$ | 27. $y'' + 5y' + 12.5y = 0$ |

Represent the following solutions in the form (5) and determine A and B such that the given conditions are satisfied.

28. $y = c_1 e^{2ix} + c_2 e^{-2ix}$, $y(0) = 0$, $y'(0) = -2$
29. $y = c_1 e^{(1+i)x} + c_2 e^{(1-i)x}$, $y(0) = 1$, $y'(0) = 1$
30. $y = c_1 e^{ix} + c_2 e^{-ix}$, $y(0) = 1$, $y'(0) = 2$
31. $y = c_1 e^{(-2+3i)x} + c_2 e^{(-2-3i)x}$, $y(0) = 0$, $y'(0) = 6$
32. $y = c_1 e^{ix} + c_2 e^{-ix}$, $y(0) = 3$, $y(\pi/2) = 0$
33. $y = c_1 e^{ix} + c_2 e^{-ix}$, $y(0) = 2$, $y(0.5) = -1$

Solve the following initial value problems.

34. $y'' + 4y = 0$, $y(0) = 0$, $y'(0) = 6$
35. $y'' - 4y = 0$, $y(0) = 2$, $y'(0) = 4$
36. $y'' + 2y' + 10y = 0$, $y(0) = 0$, $y'(0) = 3$
37. $y'' + 4y' + 5y = 0$, $y(0) = 1$, $y'(0) = -3$
38. $y'' - 2y' + 5y = 0$, $y(0) = -3$, $y'(0) = 1$
39. $y'' - 3y' + 2y = 0$, $y(0) = -1$, $y'(0) = 0$

40. Consider Ex. 3 and start from another general solution, say

$$y(x) = c_1 e^{(1+3i)x} + c_2 e^{(1-3i)x}.$$

Show that from the initial conditions we obtain $c_1 = 2 + i/2$, $c_2 = 2 - i/2$ and the resulting particular solution may be simplified so that it becomes identical with (7), which illustrates that the choice of the general solution is immaterial.

2.5 Double Root of the Characteristic Equation

To complete our discussion of the solutions of homogeneous linear differential equations with constant coefficients

$$(1) \quad y'' + ay' + by = 0$$

we shall now consider the case when the characteristic equation

$$(2) \quad \lambda^2 + a\lambda + b = 0$$

has a double root. This case is sometimes called the **critical case**. It arises if, and only if, the discriminant of (2) is zero:

$$\alpha^2 - 4b = 0, \quad \text{and then} \quad b = \frac{1}{4}\alpha^2.$$

The root is $\lambda = -a/2$, and we obtain at first only one solution

$$(3) \quad y_1 = e^{\lambda x} \quad \left(\lambda = -\frac{a}{2} \right).$$

To find another solution $y_2(x)$ we may apply the method of variation of parameters (cf. Sec. 1.8), starting from

$$y_2(x) = u(x)y_1(x) \quad \text{where} \quad y_1(x) = e^{-ax/2}.$$

In the case of a double root, (1) has the form

$$y'' + ay' + \frac{1}{4}a^2y = 0.$$

Substituting y_2 in this equation and collecting terms, we have

$$u(y_1'' + ay_1' + \frac{1}{4}a^2y_1) + u'(2y_1' + ay_1) + u''y_1 = 0.$$

Since y_1 is a solution, the expression in the first parentheses is zero. Since

$$2y_1' = 2\left(-\frac{a}{2}\right)e^{-ax/2} = -ay_1,$$

the expression in the last parentheses is zero, and our equation reduces to $u''y_1 = 0$. Hence $u'' = 0$. A solution is $u = x$. Consequently,

$$(4) \quad y_2(x) = xe^{ax} \quad \left(\lambda = -\frac{a}{2}\right).$$

We see that y_1 and y_2 are linearly independent, and our result may be formulated as follows.

Theorem 1 (Double root). *In the case of a double root of (2) the functions (3) and (4) are solutions of (1). They constitute a fundamental system. The corresponding general solution is*

$$(5) \quad y = \underline{(c_1 + c_2x)e^{ax}} \quad \left(\lambda = -\frac{a}{2}\right).$$

Note that if λ is a simple root of (2), then (4) is not a solution of (1).

Example 1. Solve

$$y'' + 8y' + 16y = 0.$$

The characteristic equation has the double root $\lambda = -4$. Hence a fundamental system of solutions is

$$e^{-4x} \quad \text{and} \quad xe^{-4x}$$

and the corresponding general solution is

$$y = (c_1 + c_2x)e^{-4x}.$$

Example 2. Solve the initial value problem

$$y'' - 4y' + 4y = 0, \quad y(0) = 3, \quad y'(0) = 1.$$

A general solution of the differential equation is

$$y(x) = (c_1 + c_2x)e^{2x}.$$

By differentiation we obtain

$$y'(x) = c_2e^{2x} + 2(c_1 + c_2x)e^{2x}.$$

From this and the initial conditions it follows that

$$y(0) = c_1 = 3, \quad y'(0) = c_2 + 2c_1 = 1.$$

Hence $c_1 = 3$, $c_2 = -5$, and the answer is

$$y = (3 - 5x)e^{2x}.$$

This completes the discussion of all three cases, and we may sum up our result in tabular form:

Case	Roots of (2)	Fundamental system of (1)	General solution of (1)
I	Distinct real λ_1, λ_2	$e^{\lambda_1 x}, e^{\lambda_2 x}$	$y = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}$
II	Complex conjugate $\lambda_1 = p + iq, \lambda_2 = p - iq$	$e^{px} \cos qx, e^{px} \sin qx$	$y = e^{px}(A \cos qx + B \sin qx)$
III	Real double root $\lambda (= -a/2)$	$e^{\lambda x}, xe^{\lambda x}$	$y = (c_1 + c_2 x)e^{\lambda x}$

Problems

Verify that the given function is a solution of the corresponding equation.

1. $y = xe^x, y' - 2y' + y = 0$
2. $y = xe^{-2x}, y'' + 6y' + 9y = 0$
3. $y = 2xe^{4x}, y'' - 8y' + 16y = 0$
4. $y = (4x - 7)e^{-x}, y'' + 2y' + y = 0$
5. $y = (c_1 x + c_2)e^{x/2}, y'' - y' + 0.25y = 0$

6. Verify directly that in the case of a double root of (2) the function (4) is a solution of (1).
7. Verify that if (2) does not have a double root, then (4) is *not* a solution of (1).

Find a general solution of the following differential equations.

- | | |
|----------------------------------|------------------------------|
| 8. $y'' + 4y' + 4y = 0$ | 9. $y'' - 0.2y' + 0.01y = 0$ |
| 10. $y'' - 2\sqrt{2}y' + 2y = 0$ | 11. $y'' + 3y' + 2.25y = 0$ |
| 12. $y'' - 4ny' + 4n^2y = 0$ | 13. $y'' - 5y' + 6.25y = 0$ |
14. Apply the method of variation of parameters to $y'' - 2y' + y = 0$, using $y_1 = e^x$.

Solve the following initial value problems.

15. $y'' - 2y' + y = 0, y(0) = 0, y'(0) = 3$
16. $y'' + 4y' + 4y = 0, y(0) = -1, y'(0) = 3$
17. $y'' - 6y' + 9y = 0, y(0) = 2, y'(0) = 8$
18. $y'' - 4y' + 4y = 0, y(0) = 4, y'(0) = 8$
19. $y'' + y' + 0.25y = 0, y(0) = -2, y'(0) = 1$
20. $y'' + 2y' + y = 0, y(0) = 1, y'(0) = -2$

2.6 Free Oscillations

We shall now consider an important engineering application of linear homogeneous differential equations with constant coefficients. The problem will

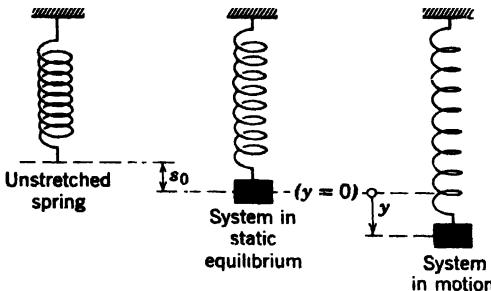


Fig. 59. Mechanical system under consideration.

be taken from mechanics, but we shall see later that it has a complete analogue in electric circuits.

Consider an ordinary spring, which resists compression as well as extension. Suppose that the spring is suspended vertically from a fixed support (Fig. 59). At the lower end of the spring we attach a body whose mass m is so large in comparison with the mass of the spring that we may disregard the latter. If the body is pulled down a certain distance and then released it will undergo a motion. We assume that the motion takes place strictly in a vertical direction.

We want to determine the motion of our mechanical system, that is, the displacement of the body as a function of the time. For this purpose we shall consider the forces⁶ acting on the body during the motion. This will lead to a differential equation, and by solving this equation we shall then obtain the displacement as a function of the time.

We choose the downward direction as the positive direction and regard forces which act downward as positive and upward forces as negative.

The most obvious force acting on the body is the *attraction of gravity*

$$(1) \quad F_1 = mg$$

where m is the mass of the body and g ($= 980 \text{ cm/sec}^2$) is the acceleration of gravity.

⁶ The most important systems of units are as follows.

System of units	Length	Mass	Time	Force
Cgs system	centimeter (cm)	gram (gm)	second (sec)	dyne
Mks system	meter	kilogram (kg)	second (sec)	newton (nt)
Engineering system	foot (ft)	slug	second (sec)	pound (lb)

$$1 \text{ ft} = 30.4800 \text{ cm} = 0.304800 \text{ meter}, \quad 1 \text{ slug} = 14,594 \text{ g} = 14.594 \text{ kg}.$$

$$1 \text{ lb} = 444,822 \text{ dynes} = 4.44822 \text{ nt}$$

For further details see, e.g., D. Halliday and R. Resnick, *Physics*. New York: Wiley, 1966.

The next force is to be considered is the *spring force* exerted by the spring, if the spring is stretched. Experiments show that this force is proportional to the stretch, say,

$$F = ks \quad (\text{Hooke's law})$$

where s is the stretch; the constant of proportionality k is called the *spring modulus*.

If $s = 1$, then $F = k$. The stiffer the spring, the larger k is.

When the body is at rest we describe its position as the *static equilibrium position*. Clearly in this position the spring is stretched by an amount s_0 such that the resultant of the corresponding spring force and the gravitational force (1) is zero. Hence this spring force acts upward and its magnitude ks_0 is equal to that of F_1 ; that is,

$$(2) \quad ks_0 = mg.$$

Let $y = y(t)$ denote the displacement of the body from the static equilibrium position, with the positive direction downward (Fig. 59).

From Hooke's law it follows that the spring force corresponding to a displacement y is

$$(3) \quad F_2 = -ks_0 - ky,$$

the resultant of the spring force $-ks_0$ when the body is in static equilibrium position and the additional spring force $-ky$ caused by the displacement. Note that the sign of the last term in (3) is properly chosen, because when y is positive, $-ky$ is negative and, according to our agreement, represents an upward force, while for negative y the force $-ky$ represents a downward force.

The resultant of the forces F_1 and F_2 given by (1) and (3) is

$$F_1 + F_2 = mg - ks_0 - ky,$$

and because of (2) this becomes

$$(4) \quad F_1 + F_2 = -ky.$$

Undamped system. If the damping of the system is so small that it can be disregarded, then (4) is the resultant of all the forces acting on the body. The differential equation will now be obtained by the use of Newton's second law

$$\text{Mass} \times \text{Acceleration} = \text{Force}$$

where *force* means the resultant of the forces acting on the body at any instant. In our case, the acceleration is $\ddot{y} = d^2y/dt^2$ and that resultant is given by (4). Thus

$$m\ddot{y} = -ky.$$

Hence the motion of our system is governed by the linear equation with constant coefficients

$$(5) \quad m\ddot{y} + ky = 0.$$

⁷ ROBERT HOOKE (1635–1703), English physicist.

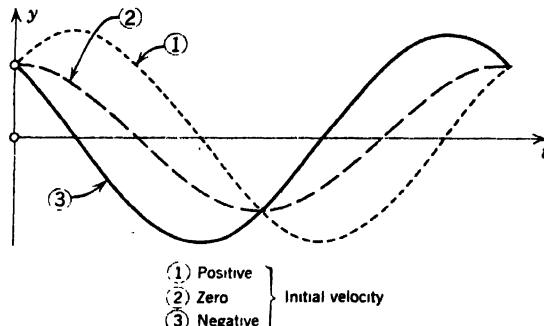


Fig. 60. Harmonic oscillations.

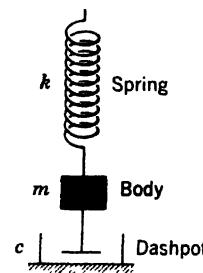


Fig. 61. Damped system.

The general solution is

$$(6) \quad y(t) = A \cos \omega_0 t + B \sin \omega_0 t \quad \text{where } \omega_0 = \sqrt{k/m}.$$

We know that this may be written (cf. Sec. 0.1)

$$(6^*) \quad y(t) = C \cos (\omega_0 t - \delta) \quad \left(\tan \delta = \frac{B}{A} \right).$$

The corresponding motion is called a **harmonic oscillation**. Figure 60 shows typical forms of (6) corresponding to various initial conditions.

Since the period of the trigonometric functions in (6) is $2\pi/\omega_0$, the body executes $\omega_0/2\pi$ cycles per second. The quantity $\omega_0/2\pi$ is called the **frequency** of the oscillations. The larger k is, the higher is the frequency, in agreement with our intuition that the stiffer the spring is, the more rapidly the body oscillates.

Damped system. If we connect the mass to a dashpot (Fig. 61), then we have to take the corresponding viscous damping into account. The corresponding damping force has the direction opposite to the instantaneous motion, and we shall assume that it is proportional to the velocity $\dot{y} = dy/dt$ of the body. This is generally a good approximation, at least for small velocities. Thus the damping force is of the form

$$F_3 = -c\dot{y}.$$

Let us show that the **damping constant** c is positive. If \dot{y} is positive, the body moves downward (in the positive y -direction) and $-c\dot{y}$ must be an upward force, that is, by agreement, $-c\dot{y} < 0$, which implies $c > 0$. For negative \dot{y} the body moves upward and $-c\dot{y}$ must represent a downward force, that is, $-c\dot{y} > 0$ which implies $c > 0$.

The resultant of the forces acting on the body is now [cf. (4)]

$$F_1 + F_2 + F_3 = -ky - c\dot{y}.$$

Hence, by Newton's second law,

$$m\ddot{y} = -ky - cy,$$

and we see that the motion of the damped mechanical system is governed by the linear differential equation with constant coefficients

$$(7) \quad m\ddot{y} + cy + ky = 0.$$

The corresponding characteristic equation is

$$\lambda^2 + \frac{c}{m}\lambda + \frac{k}{m} = 0.$$

The roots are

$$\lambda_{1,2} = -\frac{c}{2m} \pm \frac{1}{2m} \sqrt{c^2 - 4mk}.$$

Using the abbreviated notations

$$(8) \quad \alpha = \frac{c}{2m} \quad \text{and} \quad \beta = \frac{1}{2m} \sqrt{c^2 - 4mk},$$

we can write

$$\lambda_1 = -\alpha + \beta \quad \text{and} \quad \lambda_2 = -\alpha - \beta.$$

The form of the solution of (7) will depend on the damping, and, as in Sec. 2.2, we may now distinguish between the following three cases:

Case I. $c^2 > 4mk$. Distinct real roots λ_1, λ_2 . (Overdamping.)

Case II. $c^2 < 4mk$. Complex conjugate roots. (Underdamping.)

Case III. $c^2 = 4mk$. A real double root. (Critical damping.)

Let us discuss these three cases separately.

Case I. Overdamping. When the damping constant c is so large that $c^2 > 4mk$, then λ_1 and λ_2 are distinct real roots, the general solution of (7) is

$$(9) \quad y(t) = c_1 e^{-(\alpha-\beta)t} + c_2 e^{-(\alpha+\beta)t},$$

and we see that in this case the body does not oscillate. For $t > 0$ both exponents in (9) are negative because $\alpha > 0$, $\beta > 0$, and $\beta^2 = \alpha^2 - k/m < \alpha^2$. Hence both terms in (9) approach zero as t approaches infinity. Practically speaking, after a sufficiently long time the mass will be at rest in the static equilibrium position ($y = 0$). Figure 62 shows (9) for some typical initial conditions.

Case II. Underdamping. This is the most interesting case. If the damping constant c is so small that $c^2 < 4mk$, then β in (8) is purely imaginary, say

$$(10) \quad \beta = i\omega^* \quad \text{where} \quad \omega^* = \frac{1}{2m} \sqrt{4mk - c^2} \quad (>0).$$

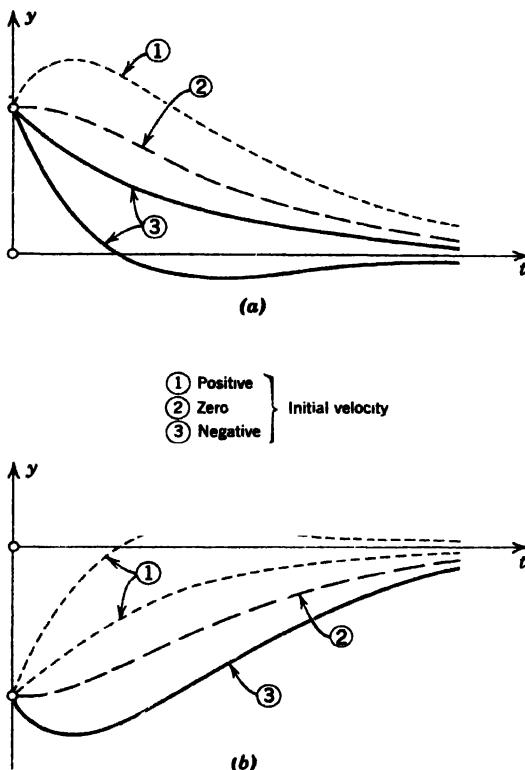


Fig. 62. Typical motions in the overdamped case. (a) Positive initial displacement.
(b) Negative initial displacement.

The roots of the characteristic equation are complex conjugate,

$$\lambda_1 = -\alpha + i\omega^*, \quad \lambda_2 = -\alpha - i\omega^*,$$

and the general solution is

$$(11) \quad y(t) = e^{-\alpha t}(A \cos \omega^* t + B \sin \omega^* t) = Ce^{-\alpha t} \cos(\omega^* t - \delta)$$

where $C^2 = A^2 + B^2$ and $\tan \delta = B/A$ (cf. Sec. 0.1).

This solution represents damped oscillations. Since $\cos(\omega^* t - \delta)$ varies between -1 and 1 , the curve of the solution lies between the curves $y = Ce^{-\alpha t}$ and $y = -Ce^{-\alpha t}$ in Fig. 63, touching these curves when $\omega^* t - \delta$ is an integral multiple of π .

The frequency is $\omega^*/2\pi$ cycles per second. From (10) we see that the smaller $c (> 0)$ is, the larger is ω^* and the more rapid the oscillations become. As c approaches zero, ω^* approaches the value $\omega_0 = \sqrt{k/m}$, corresponding to the harmonic oscillation (6).

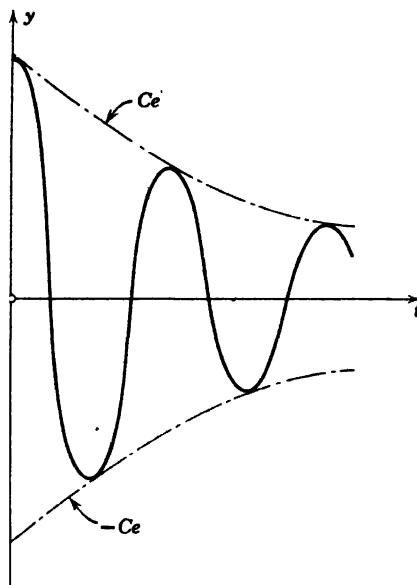


Fig. 63. Damped oscillation in Case II.

Case III. Critical damping. If $c^2 = 4mk$, then $\beta = 0$, $\lambda_1 = \lambda_2 = -\alpha$, and the general solution is

$$(12) \quad y(t) = (c_1 t + c_2)e^{-\alpha t}.$$

Since the exponential function is never zero and $c_1 t + c_2$ can have at most one positive zero, it follows that the motion can have at most one passage through the equilibrium position ($y = 0$). If the initial conditions are such that c_1 and c_2 have the same sign, there is no such passage at all. This is similar to Case I. Figure 64 shows typical forms of (12).

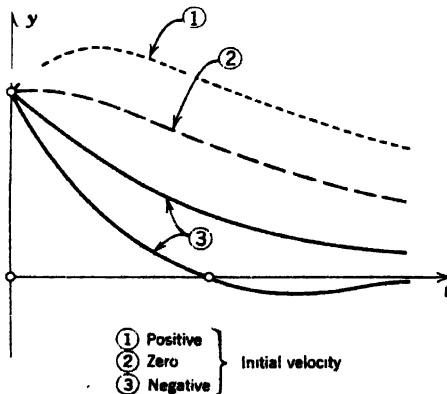


Fig. 64. Critical damping.

Problems

1. Consider the harmonic oscillation (6) with $\omega_0 = 1$. Determine A and B such that the initial displacement is 1 in. and the initial velocity is (a) -1 in./sec, (b) 0, (c) 1 in./sec. Represent these three functions in the form (6*) and graph their curves. What are the displacements and velocities when $t = \pi$ and $t = 2\pi$?
2. Using (6), find the harmonic oscillation having the initial displacement 1 and the initial velocity 0. Graph the curves for $m = 1$ and various values of k , say $k = 1, 4, 9$, and 16. What are the frequencies of these oscillations?
3. Find the harmonic oscillation starting with initial displacement 1 and initial velocity 1. Represent the oscillation in the form (6*). Graph the curves for $m = 1$ and $k = 1, 4, 9$, and 16.
4. Find the harmonic oscillation (6) starting with initial displacement y_0 and initial velocity v_0 and represent it in the form (6*).
5. Does the frequency of a harmonic or damped oscillation depend on the initial conditions?
6. Determine the motions of the mechanical system described by (7), starting from $y = 1$ with initial velocity 0. Choose $m = 1$, $k = 1$, and increasing values of the damping constant, say $c = 0, 1, 2, 10$. Graph the solutions $y(t)$.
7. Find the overdamped motion (9) starting with initial displacement y_0 and initial velocity v_0 .
8. Prove that in the overdamped case the mass cannot pass through $y = 0$ more than once (Fig. 62).
9. Prove that if the initial conditions are such that both constants c_1 and c_2 in (9) have the same sign, then the mass does not pass through $y = 0$.
10. Prove that if in (9) the mass starts from $y = 0$, it cannot pass through $y = 0$ again.
11. How does the frequency of a damped oscillation depend on the damping?
12. Find the underdamped motion (11) starting with initial displacement y_0 and initial velocity v_0 .
13. Determine A and B such that the damped oscillation
$$y = e^{-t}(A \cos t + B \sin t)$$
starts with initial displacement 1 and initial velocity (a) -1, (b) 0, (c) 1. Represent these three functions in the form $y = Ce^{-t} \cos(t - \delta)$, graph them, and compare the motions with those in Prob. 1.
14. How does the position of the maximum and minimum displacements of the oscillation $y = e^{-at} \cos t$ depend on the damping?
15. Prove that the maxima and minima of a damped oscillation of the form (11) occur at equidistant values of t , the distance between two consecutive maxima being $2\pi/\omega^*$.
16. (Logarithmic decrement) Prove that the ratio of two consecutive maximum amplitudes of a damped oscillation (11) is constant, the natural logarithm of this ratio being $\Delta = 2\pi\alpha/\omega^*$. (Δ is called the *logarithmic decrement* of the oscillation.)
17. Find the values of t at which $y = e^{-t} \cos t$ has maxima and minima. To check your result, graph $y(t)$. Find the logarithmic decrement.
18. Find the critical motion (12) which starts from y_0 with initial velocity v_0 .
19. Under what conditions does (12) have a maximum or minimum at some instant $t > 0$?
20. Represent the maximum or minimum amplitude in Prob. 19 in terms of the initial values y_0 and v_0 .

2.7 Cauchy Equation

The so-called **Cauchy equation** or **Euler equation**

$$(1) \quad x^2y'' + axy' + by = 0, \quad (a, b \text{ constant})$$

can also be solved by purely algebraic manipulations. By substituting

$$(2) \quad y = x^m.$$

and its derivatives into the equation (1) we find

$$x^2m(m-1)x^{m-2} + axmx^{m-1} + bx^m = 0.$$

By omitting the common power x^m , which is not zero when $x \neq 0$, we obtain the auxiliary equation

$$(3) \quad m^2 + (a-1)m + b = 0.$$

If the roots m_1 and m_2 of this equation are different, then the functions

$$y_1(x) = x^{m_1} \quad \text{and} \quad y_2(x) = x^{m_2}$$

constitute a fundamental system of solutions of (1) for all x for which these functions are defined. The corresponding general solution is

$$(4) \quad y = c_1x^{m_1} + c_2x^{m_2} \quad (c_1, c_2 \text{ arbitrary}).$$

Example 1. Solve

$$x^2y'' - 1.5xy' - 1.5y = 0.$$

The auxiliary equation is

$$m^2 - 2.5m - 1.5 = 0.$$

The roots are $m_1 = -0.5$ and $m_2 = 3$. Hence a fundamental system of real solutions for all positive x is

$$y_1 = \frac{1}{\sqrt{x}}, \quad y_2 = x^3$$

and the corresponding general solution for all those x is

$$y = \frac{c_1}{\sqrt{x}} + c_2x^3.$$

The auxiliary equation (3) has a *double root* $m_1 = m_2$ if, and only if,

$$b = \frac{1}{4}(a-1)^2, \quad \text{and then} \quad m_1 = m_2 = \frac{1-a}{2}.$$

In this *critical case* we may obtain a second solution by applying the method of variation of parameters. The procedure is similar to that in Sec. 2.5 and yields (cf. Prob. 1)

$$y_2 = u y_1 = (\ln x)y_1.$$

Thus, writing m in place of m_1 , we see that

$$(5) \quad y_1 = x^m \quad \text{and} \quad y_2 = x^m \ln x \quad \left(m = \frac{1-a}{2} \right)$$

are solutions of (1) in the case of a double root m of (3). Since these solutions are linearly independent, they constitute a fundamental system of real solutions of (1) for all positive x , and the corresponding general solution is

$$(6) \quad y = (c_1 + c_2 \ln x)x^m, \quad (c_1, c_2 \text{ arbitrary}).$$

Example 2. Solve

$$x^2y'' - 3xy' + 4y = 0.$$

The auxiliary equation has the double root $m = 2$. Hence a fundamental system of real solutions for all positive x is $x^2, x^2 \ln x$, and the corresponding general solution is

$$y = (c_1 + c_2 \ln x)x^2.$$

Problems

1. Derive y_2 in (5) by the method of variation of parameters.
2. Verify directly that in the case of a double root of (3) the function y_2 in (5) is a solution of (1).
3. Show that if the roots m_1 and m_2 of (3) are different, the functions $x^{m_1} \ln x$ and $x^{m_2} \ln x$ are *not* solutions of (1).

Find a general solution of the following differential equations.

- | | |
|-------------------------------|-------------------------------|
| 4. $x^2y'' + xy' - y = 0$ | 5. $x^2y'' - 3xy' + 3y = 0$ |
| 6. $x^2y'' + xy' - 4y = 0$ | 7. $x^2y'' - xy' + y = 0$ |
| 8. $x^2y'' - xy' + 0.75y = 0$ | 9. $xy'' - 3y' = 0$ |
| 10. $x^2y'' + 0.25y = 0$ | 11. $xy'' + y' = 0$ |
| 12. $x^2y'' + 5xy' + 3y = 0$ | 13. $x^2y'' + 3.5xy' + y = 0$ |

Solve the following initial value problems.

14. $x^2y'' + xy' - 0.25y = 0, \quad y(1) = 2, \quad y'(1) = 1$
15. $x^2y'' - 3xy' + 4y = 0, \quad y(1) = 1, \quad y'(1) = 1$
16. $x^2y'' + xy' - 2.25y = 0, \quad y(1) = 2, \quad y'(1) = 0$
17. $x^2y'' - 3xy' + 3y = 0, \quad y(1) = 0, \quad y'(1) = -2$
18. Show that by setting $x = e^t$ ($t > 0$) the Cauchy equation (1) can be transformed into the equation $\ddot{y} + (a-1)\dot{y} + by = 0$, whose coefficients are constant.
19. Transform the equation in Prob. 18 back into (1).
20. Show that if we apply the transformation in Prob. 18, then (2) yields an expression of the form (2), Sec. 2.2, and (6) yields an expression of the form (5), Sec. 2.5, except for notation.

2.8 Existence and Uniqueness of Solutions

An *initial value problem for a second-order differential equation* consists of the equation and two initial conditions, one for the solution $y(x)$ and one for its

derivative $y'(x)$. We shall now consider initial value problems of the form

$$(1) \quad y'' + f(x)y' + g(x)y = 0, \quad y(x_0) = K_1, \quad y'(x_0) = K_2$$

involving a second-order linear homogeneous differential equation; K_1 and K_2 are given constants. The important result will be that continuity of the coefficients f and g is sufficient for the existence and uniqueness of the solution of the problem (1).

Existence and uniqueness theorem. *If $f(x)$ and $g(x)$ are continuous functions on an open interval I and x_0 is in I , then the initial value problem (1) has a unique solution $y(x)$ on I .*

The *proof of existence* uses the same prerequisites as those of the existence theorem in Sec. 1.12 and will not be presented here; it can be found in Ref. [B5] in Appendix 1.

Proof of uniqueness.⁸ Assuming that (1) has two solutions $y_1(x)$ and $y_2(x)$ on I , we show that their difference

$$y(x) = y_1(x) - y_2(x)$$

is identically zero on I ; then $y_1 \equiv y_2$ on I , which implies uniqueness.

Since the equation in (1) is homogeneous and linear, y is a solution of that equation on I , and since y_1 and y_2 satisfy the same initial conditions, y satisfies the conditions

$$(2) \quad y(x_0) = 0, \quad y'(x_0) = 0.$$

We consider the function

$$z(x) = y(x)^2 + y'(x)^2$$

and its derivative

$$z' = 2yy' + 2y'y''.$$

From the differential equation we have

$$y'' = -fy' - gy.$$

By substituting this in the expression for z' we obtain

$$(3) \quad z' = 2yy' - 2fy'^2 - 2gyy'.$$

Now, since y and y' are real,

$$(y \pm y')^2 = y^2 \pm 2yy' + y'^2 \geq 0,$$

and from this we immediately obtain the two inequalities

$$(4) \quad (a) \quad 2yy' \leq y^2 + y'^2 = z, \quad (b) \quad -2yy' \leq y^2 + y'^2 = z.$$

For the last term in (3) we now have

$$-2gyy' \leq | -2gyy' | = |g| |2yy'| \leq |g| z.$$

⁸ This proof was suggested by my colleague, Prof. A. D. Ziebur.

Using this result and applying (4a) to the term $2yy'$ in (3), we find

$$z' \leqq z + 2|f|y'^2 + |g|z.$$

Since $y'^2 \leqq y^2 + y'^2 = z$ we obtain from this

$$z' \leqq (1 + 2|f| + |g|)z$$

or, denoting the function in parentheses by h ,

$$(5) \quad z' \leqq hz \quad \text{for all } x \text{ on } I.$$

Similarly, from (3) and (4) it follows that

$$(6) \quad \begin{aligned} -z' &= -2yy' + 2fy'^2 + 2gyy' \\ &\leqq z + 2|f|z + |g|z = hz. \end{aligned}$$

The inequalities (5) and (6) are equivalent to the inequalities

$$(7) \quad z' - hz \leqq 0, \quad z' + hz \geqq 0.$$

Integrating factors for the two expressions on the left are

$$F_1 = e^{-\int h(x) dx} \quad \text{and} \quad F_2 = e^{\int h(x) dx}.$$

The integrals in the exponents exist since h is continuous. Since F_1 and F_2 are positive, we thus obtain from (7)

$$F_1(z' - hz) = (F_1 z)' \leqq 0 \quad \text{and} \quad F_2(z' + hz) = (F_2 z)' \geqq 0,$$

which means that $F_1 z$ is nonincreasing and $F_2 z$ is nondecreasing on I . Since by (2), $z(x_0) = 0$ we thus obtain when $x \leqq x_0$

$$F_1 z \geqq (F_1 z)_{x_0} = 0, \quad F_2 z \leqq (F_2 z)_{x_0} = 0$$

and similarly, when $x \geqq x_0$,

$$F_1 z \leqq 0, \quad F_2 z \geqq 0.$$

Dividing by F_1 and F_2 and noting that these functions are positive, we altogether have

$$z \leqq 0, \quad z \geqq 0 \quad \text{for all } x \text{ on } I.$$

This implies $z = y^2 + y'^2 \equiv 0$ on I . Hence $y \equiv 0$ or $y_1 \equiv y_2$ on I , and the proof is complete.

The remaining part of the present section will be devoted to some important consequences of this theorem.

To prepare for this consideration let us return to the concepts of linear dependence and independence of functions introduced in Sec. 2.3. Let $y_1(x)$ and $y_2(x)$ be given functions defined on an interval I and consider the relation

$$k_1 y_1(x) + k_2 y_2(x) = 0$$

on I , where k_1 and k_2 are constants to be determined. This relation certainly holds for $k_1 = 0, k_2 = 0$. It may happen that we can find a $k_1 \neq 0$ such that the relation holds on I . Then we may divide by k_1 and solve for y_1 :

$$y_1 = -\frac{k_2}{k_1} y_2.$$

Similarly, if that relation holds for a $k_2 \neq 0$ on I we may solve for y_2 :

$$y_2 = -\frac{k_1}{k_2} y_1.$$

We see that in both cases y_1 and y_2 are linearly dependent on I .

However, if $k_1 = 0, k_2 = 0$ is the only pair of numbers for which that relation holds on I , we cannot solve for y_1 or y_2 , and the functions are linearly independent on I .

Consequently, two functions $y_1(x)$ and $y_2(x)$ are linearly dependent on an interval I where both functions are defined if, and only if, we can find constants k_1, k_2 , not both zero such that

$$k_1 y_1(x) + k_2 y_2(x) = 0$$

for all x on I .

This result will be used in the subsequent proof.

In our further consideration we shall need another criterion for testing linear dependence and independence of two solutions $y_1(x)$ and $y_2(x)$ of the differential equation

$$(8) \quad y'' + f(x)y' + g(x)y = 0.$$

This criterion uses the determinant

$$W(y_1, y_2) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = y_1 y_2' - y_2 y_1',$$

which is called the *Wronski determinant*⁹ or briefly the *Wronskian* of y_1 and y_2 . It can be stated as follows.

Theorem 2. Suppose that the coefficients $f(x)$ and $g(x)$ of the differential equation (8) are continuous on an open interval I . Then any two solutions of (8) on I are linearly dependent on I if, and only if, their Wronskian W is zero for some $x = x_0$ in I . (If $W = 0$ for $x = x_0$, then $W \equiv 0$ on I .)

Proof. (a) Let $y_1(x)$ and $y_2(x)$ be linearly dependent solutions of (8) on I . Then there are constants k_1 and k_2 , not both zero, such that

$$k_1 y_1 + k_2 y_2 = 0$$

holds on I . By differentiation we obtain

$$k_1 y_1' + k_2 y_2' = 0.$$

⁹ Introduced by HÖNE (1778–1853), who changed his name to Wronski.

Suppose that $k_1 \neq 0$. Then by multiplying the first equation by y_2'/k_1 and the last by $-y_2/k_1$ and adding the resulting equations, we obtain

$$y_1 y_2' - y_2 y_1' = 0, \quad \text{that is,} \quad W(y_1, y_2) = 0.$$

If $k_1 = 0$, then $k_2 \neq 0$, and the argument is similar.

(b) Conversely, suppose that $W(y_1, y_2) = 0$ for some $x = x_0$ in I . Consider the system of linear equations

$$(9) \quad \begin{aligned} k_1 y_1(x_0) + k_2 y_2(x_0) &= 0 \\ k_1 y_1'(x_0) + k_2 y_2'(x_0) &= 0 \end{aligned}$$

in the unknowns k_1, k_2 . Since the system is homogeneous and the determinant of the system is just the Wronskian $W[y_1(x_0), y_2(x_0)]$ and $W = 0$, the system has a solution k_1, k_2 where k_1 and k_2 are not both zero (cf. Sec. 0.3). Using these numbers k_1, k_2 , we introduce the function

$$y(x) = k_1 y_1(x) + k_2 y_2(x).$$

By Fundamental Theorem 1 in Sec. 2.1 the function $y(x)$ is a solution of (8) on I . From (9) we see that it satisfies the initial conditions $y(x_0) = 0$, $y'(x_0) = 0$. Now another solution of (8) satisfying the same initial conditions is $y^* \equiv 0$. Since f and g are continuous, the uniqueness theorem is applicable and we conclude that $y^* \equiv y$, that is,

$$k_1 y_1 + k_2 y_2 \equiv 0$$

on I . Since k_1 and k_2 are not both zero, this means that y_1 and y_2 are linearly dependent on I . Furthermore, from this and the first part of the proof it follows that $W(y_1, y_2) \equiv 0$ on I .

Example 1. The functions $y_1 = \cos \omega x$ and $y_2 = \sin \omega x$ are solutions of the equation

$$y'' + \omega^2 y = 0 \quad (\omega \neq 0).$$

A simple calculation shows that their Wronskian has the value ω . Hence these functions constitute a fundamental system of solutions for all x .

Further examples are included in Secs. 2.3–2.5.

The following example illustrates the fact that *the assumption of continuity of the coefficients in Theorem 2 cannot be omitted*.

Example 2. The functions

$$y_1 = x^3 \quad \text{and} \quad y_2 = \begin{cases} x^3 & \text{when } x \geq 0 \\ -x^3 & \text{when } x \leq 0 \end{cases}$$

are solutions of the differential equation

$$(10) \quad y'' - \frac{3}{x} y' + \frac{3}{x^2} y = 0$$

for all x , as can be readily verified by substituting each of the functions and its derivatives into (10). The reader may show that the functions are linearly independent on the x -axis, although their Wronskian is identically zero. This does not contradict Theorem 2, because the coefficients of (10) are not continuous at $x = 0$.

Of course, Ex. 2 also illustrates the fact that if u_1 and u_2 are any differentiable functions of x , the condition $W(u_1, u_2) = 0$ is only necessary for linear dependence of u_1 and u_2 (as follows from part (a) of the proof of Theorem 2), but not sufficient.

Let us now show that an equation of the form (8), whose coefficients are continuous on an open interval I , has a fundamental system of solutions on I , and, in fact, infinitely many such systems.

Proof. From the existence theorem it follows that (8) has a solution $y_1(x)$ satisfying the initial conditions

$$y_1(x_0) = 1, \quad y_1'(x_0) = 0,$$

and a solution $y_2(x)$ satisfying the initial conditions

$$y_2(x_0) = 0, \quad y_2'(x_0) = 1,$$

where x_0 is any fixed point in I . At x_0 , $W(y_1, y_2) = 1$. Hence, by Theorem 2, the two solutions are linearly independent on I and form a fundamental system of (8) on I . Another fundamental system is ay_1, by_2 where $a \neq 0, b \neq 0$ are any constants, so that there are infinitely many such systems. This completes the proof.

We shall now prove the important fact that a general solution of an equation of the form (8) with continuous coefficients includes the totality of all solutions of (8).

Theorem 3. Let

$$(11) \quad y(x) = c_1 y_1(x) + c_2 y_2(x)$$

be a general solution of the differential equation (8) on an open interval I where the coefficients f and g are continuous. Let $Y(x)$ be any solution of (8) on I containing no arbitrary constants. Then $Y(x)$ is obtained from (11) by assigning suitable values to the constants c_1 and c_2 .

Proof. Let $Y(x)$ be any solution of (8) on I containing no arbitrary constants, and let $x = x_0$ be any fixed point in I . We first show that we can find values of c_1 and c_2 in (11) such that

$$y(x_0) = Y(x_0), \quad y'(x_0) = Y'(x_0)$$

or at length

$$(12) \quad \begin{aligned} c_1 y_1(x_0) + c_2 y_2(x_0) &= Y(x_0), \\ c_1 y_1'(x_0) + c_2 y_2'(x_0) &= Y'(x_0). \end{aligned}$$

In fact, this is a system of linear equations in the unknowns c_1, c_2 . Its determinant is the Wronskian of y_1 and y_2 at $x = x_0$. Since (11) is a general solution, y_1 and y_2 are linearly independent on I , and from Theorem 2 it follows that their Wronskian is not zero. Hence the system has a unique solution $c_1 = c_1^*, c_2 = c_2^*$ which can be obtained by Cramer's rule (Sec. 0.3).

By using these constants we obtain from (11) the particular solution

$$y^*(x) = c_1 * y_1(x) + c_2 * y_2(x).$$

From (12) we see that

$$(13) \quad y^*(x_0) = Y(x_0), \quad y^{*\prime}(x_0) = Y'(x_0).$$

From this and the uniqueness theorem we conclude that y^* and Y must be identical on I , and the proof is complete.

Problems

Find the Wronskian of the following functions.

1. $y_1 = x^{m_1}$, $y_2 = x^{m_2}$ 2. $y_1 = e^x \cos 2x$, $y_2 = e^x \sin 2x$

3. $y_1 = x^m$, $y_2 = x^m \ln x$ 4. $y_1 = e^{\lambda_1 x}$, $y_2 = e^{\lambda_2 x}$

5. $y_1 = e^{\lambda x}$, $y_2 = xe^{\lambda x}$ 6. $y_1 = e^{(p+iq)x}$, $y_2 = e^{(p-iq)x}$

7. Show that y_1 and y_2 in Ex. 2 constitute a fundamental system of solutions of (10), but $W(y_1, y_2) \equiv 0$.

8. Let y_1, y_2 be a fundamental system of solutions of a differential equation satisfying the assumptions of Theorem 2. Under what conditions will two linear combinations $z_1 = a_{11}y_1 + a_{12}y_2$, $z_2 = a_{21}y_1 + a_{22}y_2$ (a_{11} , etc., constant) constitute a fundamental system of that equation?

In each case find a second-order linear homogeneous differential equation for which the given functions are solutions. Using Theorem 2, show that the functions are linearly independent.

9. $\cos x$, $\sin x$ 10. e^x , e^{2x} 11. e^x , xe^x 12. x^2 , $x^{1/2}$

13. 1 , x^2 14. 1 , e^{-x} 15. x^2 , $x^2 \ln x$ 16. $e^x \cos x$, $e^x \sin x$

2.9 Homogeneous Linear Equations of Arbitrary Order

A differential equation of n th order is said to be **linear** if it can be written in the form

$$(1) \quad y^{(n)} + f_{n-1}(x)y^{(n-1)} + \cdots + f_1(x)y' + f_0(x)y = r(x)$$

where the function r on the right-hand side and the coefficients f_0, f_1, \dots, f_{n-1} are any given functions of x , and $y^{(n)}$ is the n th derivative of y with respect to x , etc. Any differential equation of order n which cannot be written in the form (1) is said to be **nonlinear**.

If $r(x) \equiv 0$, equation (1) becomes

$$(2) \quad y^{(n)} + f_{n-1}(x)y^{(n-1)} + \cdots + f_1(x)y' + f_0(x)y = 0$$

and is said to be **homogeneous**. If $r(x)$ is not identically zero, the equation is said to be **nonhomogeneous**.

A function $y = \phi(x)$ is called a **solution** of a (linear or nonlinear) differential equation of n th order on an interval I , if $\phi(x)$ is defined and n times differentiable on I and is such that the equation becomes an identity when we replace the unspecified function y and its derivatives in the equation by ϕ and its corresponding derivatives.

We shall now see that our considerations in Secs. 2.3 and 2.8 can be readily extended to linear homogeneous equations of arbitrary order n .

Existence and uniqueness theorem. If $f_0(x), \dots, f_{n-1}(x)$ in (2) are continuous functions on an open interval I , then the initial value problem consisting of the equation (2) and the n initial conditions

$$y(x_0) = K_1, \quad y'(x_0) = K_2, \quad \dots, \quad y^{(n-1)}(x_0) = K_n$$

has a unique solution $y(x)$ on I ; here x_0 is any fixed point in I , and K_1, \dots, K_n are given numbers.

The proof of existence can be found in Ref. [B5] in Appendix 1, and uniqueness can be proved by a slight generalization of the proof in Sec. 2.8.

To extend the considerations of Sec. 2.3 we need the concepts of linear dependence and independence of n functions, say, $y_1(x), \dots, y_n(x)$. These functions are said to be **linearly dependent on some interval I** where they are defined, if (at least) one of them can be represented on I as a sum of the other functions, each multiplied by some constant (zero or not). If none of the functions can be represented in this way, they are said to be **linearly independent on I** .

This definition includes that in Sec. 2.3 for two functions as a particular case.

For example, the functions $y_1 = x$, $y_2 = x^2$, $y_3 = x^3$ are linearly independent on any interval. The functions $y_1 = x$, $y_2 = 3x$, $y_3 = x^2$ are linearly dependent on any interval, because $y_2 = 3y_1 + 0y_3$.

The n functions $y_1(x), \dots, y_n(x)$ are linearly dependent on an interval I where they are defined if, and only if, we can find constants k_1, \dots, k_n , not all zero, such that the relation

$$(3) \quad k_1 y_1(x) + \dots + k_n y_n(x) = 0$$

holds for all x on I .

In fact, if (3) holds for some $k_1 \neq 0$, we may divide by k_1 and express y_1 as a linear function of y_2, \dots, y_n :

$$y_1 = -\frac{1}{k_1} (k_2 y_2 + \dots + k_n y_n),$$

which proves linear dependence. Similarly, if (3) holds for some $k_i \neq 0$, we may express y_i as a linear function of the other functions. On the other hand, if $k_1 = 0, \dots, k_n = 0$ is the only set of constants for which (3) holds

on I , then we cannot solve for any of the functions in terms of the others, and the functions are linearly independent. This proves the statement.

A solution of a differential equation of the order n (linear or not) is called a **general solution** if it contains n arbitrary¹⁰ independent constants. Here *independence* means that the same solution cannot be reduced to a form containing less than n arbitrary constants. If we assign definite values to the n constants, then the solution so obtained is called a **particular solution** of that equation.

A set of n linearly independent solutions $y_1(x), \dots, y_n(x)$ of the linear homogeneous equation (2) on I is called a **fundamental system** of solutions of (2) on I .

If y_1, \dots, y_n is such a fundamental system, then

$$(4) \quad y(x) = c_1 y_1(x) + \dots + c_n y_n(x) \quad (c_1, \dots, c_n \text{ arbitrary})$$

is a general solution of (2) on I . In fact, since (2) is linear and homogeneous, (4) is a solution of (2), and since y_1, \dots, y_n are linearly independent functions on I , none of them can be expressed as a linear function of the others, that is, (4) cannot be reduced to a form containing less than n arbitrary constants.

The test for linear dependence and independence of solutions (Theorem 2 in Sec. 2.8) can be generalized to n th order equations as follows.

Theorem 2. Suppose that the coefficients $f_0(x), \dots, f_{n-1}(x)$ of (2) are continuous on an open interval I . Then n solutions y_1, \dots, y_n of (2) on I are linearly dependent on I if, and only if, their **Wronskian**

$$(5) \quad W(y_1, \dots, y_n) = \begin{vmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & \cdots & y_n' \\ \cdot & \cdot & \cdots & \cdot \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{vmatrix}$$

is zero for some $x = x_0$ in I . (If $W = 0$ at $x = x_0$, then $W \equiv 0$ on I .)

The idea of the proof is similar to that of Theorem 2 in Sec. 2.8.

A general solution of an equation of the form (2) with continuous coefficients includes the totality of all solutions of (2). In fact, the following generalization of Theorem 3 in Sec. 2.8 holds true.

Theorem 3. Let (4) be a general solution of (2) on an open interval I where $f_0(x), \dots, f_{n-1}(x)$ are continuous, and let $Y(x)$ be any solution of (2) on I involving no arbitrary constants. Then $Y(x)$ is obtained from (4) by assigning suitable values to the arbitrary constants c_1, \dots, c_n .

The idea of the proof is similar to that of Theorem 3 in Sec. 2.8.

¹⁰ Cf. footnote 2, in Sec. 2.3.

Example 1. The functions $y_1 = e^{-x}$, $y_2 = e^x$, and $y_3 = e^{2x}$ are solutions of the equation

$$(6) \quad y''' - 2y'' - y' + 2y = 0.$$

The Wronskian is

$$W(e^{-x}, e^x, e^{2x}) = \begin{vmatrix} e^{-x} & e^x & e^{2x} \\ -e^{-x} & e^x & 2e^{2x} \\ e^{-x} & e^x & 4e^{2x} \end{vmatrix} = 6e^{3x} \neq 0,$$

which shows that the functions constitute a fundamental system of solutions of (6) for all x , and the corresponding general solution is

$$y = c_1 e^{-x} + c_2 e^x + c_3 e^{2x}.$$

Problems

Prove

1. If $y \equiv 0$ is a function of a set of functions on an interval I , the set is linearly dependent on I .
2. If one of the functions of a set of functions on I is a constant multiple of another function of the set on I , the set is linearly dependent on I .
3. If a set of p functions on I is linearly dependent on I , a set of n ($\geq p$) functions on I containing that set is linearly dependent on I .

Are the following functions linearly dependent or independent on the x -axis?

- | | | |
|------------------------------|------------------------------|--------------------------|
| 4. $1, x, x^2$ | 5. $\cosh^2 x, \sinh^2 x, 1$ | 6. $\cos x, \sin x, 1$ |
| 7. $\ln x, \ln x^2, \ln x^3$ | 8. e^x, xe^x, x^2e^x | 9. $1, x, 2x$ |
| 10. $x, x - 1, x + 1$ | 11. $1, x^2, x^4, x^6$ | 12. $e^{2x}, e^{-2x}, 0$ |

Using Theorem 2, show that the given functions form a fundamental system of solutions of the corresponding differential equation.

13. $e^x, e^{2x}, e^{3x}, y'' - 6y'' + 11y' - 6y = 0$
14. $1, e^{2x}, e^{-2x}, y''' - 4y' = 0$
15. $e^{-x}, e^x, xe^x, y''' - y'' - y' + y = 0$
16. $e^{-x}, xe^{-x}, x^2e^{-x}, y''' + 3y'' + 3y' + y = 0$
17. $x, x^2, x^3, x^3y''' - 3x^2y'' + 6xy' - 6y = 0$
18. $x^{1/2}, x^{-1/2}, 1, x^2y''' + 3xy'' + 0.75y' = 0$
19. $e^x, xe^x, x^2e^x, x^3e^x, y^{IV} - 4y''' + 6y'' - 4y' + y = 0$
20. \dots

2.10 Homogeneous Linear Equations of Arbitrary Order with Constant Coefficients

The method in Sec. 2.2 may easily be extended to a linear homogeneous equation of any order n with constant coefficients

$$(1) \quad y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = 0.$$

Substituting $y = e^{\lambda x}$ and its derivatives into (1), we obtain the *characteristic equation*

$$(2) \quad \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0.$$

If this equation has n distinct roots $\lambda_1, \dots, \lambda_n$, then the n solutions

$$(3) \quad y_1 = e^{\lambda_1 x}, \quad \dots, \quad y_n = e^{\lambda_n x}$$

constitute a fundamental system for all x , and the corresponding general solution of (1) is

$$(4) \quad y = c_1 e^{\lambda_1 x} + \cdots + c_n e^{\lambda_n x}.$$

If a *double root* occurs, say, $\lambda_1 = \lambda_2$, then $y_1 = y_2$ in (3) and we take y_1 and xy_1 as two linearly independent solutions corresponding to that root; this is as in Sec. 2.5.

If a *triple root* occurs, say, $\lambda_1 = \lambda_2 = \lambda_3$, then $y_1 = y_2 = y_3$ in (3) and three linearly independent solutions corresponding to that root are

$$(5) \quad y_1, \quad xy_1, \quad x^2y_1;$$

these solutions are obtained by the method of variation of parameters. More generally, if λ is a *root of order m* , then

$$e^{\lambda x}, \quad xe^{\lambda x}, \quad \dots, \quad x^{m-1}e^{\lambda x}$$

are m linearly independent solutions of (1) corresponding to that root.

Let us illustrate these simple facts by two examples.

Example 1. Solve the differential equation

$$y''' - 2y'' - y' + 2y = 0.$$

The roots of the characteristic equation

$$\lambda^3 - 2\lambda^2 - \lambda + 2 = 0$$

are -1 , 1 , and 2 , and the corresponding general solution (4) is

$$y = c_1 e^{-x} + c_2 e^x + c_3 e^{2x}.$$

Example 2. Solve the differential equation

$$y^{(5)} - 3y^{(4)} + 3y''' - y'' = 0.$$

The characteristic equation

$$\lambda^5 - 3\lambda^4 + 3\lambda^3 - \lambda^2 = 0$$

has the roots $\lambda_1 = \lambda_2 = 0$ and $\lambda_3 = \lambda_4 = \lambda_5 = 1$, and the answer is

$$(6) \quad y = c_1 + c_2 x + (c_3 + c_4 x + c_5 x^3)e^x.$$

Problems

- Show that (6) may also be obtained by setting $y'' = z$, finding a general solution z of the new third-order equation, and integrating z twice.
- Using the method of variation of parameters, find a general solution of

$$y''' - 3y'' + 3y' - y = 0.$$

3. Applying the method of variation of parameters to the case of a triple root of (2) ($n = 3$), obtain (5).

Find a general solution of the following differential equations.

$$\begin{array}{lll} 4. y^{IV} - y = 0 & 5. y''' + y' = 0 & 6. y'' - 4y' = 0 \\ 7. y^{IV} - 5y'' + 4y = 0 & 8. y^{IV} + 3y'' - 4y = 0 & 9. y^{IV} + 5y'' + 4y = 0 \end{array}$$

10. The Cauchy equation or Euler equation of the third order is

$$x^3y''' + ax^2y'' + bxy' + cy = 0.$$

Show that $y = x^m$ is a solution of the equation if, and only if, r is a root of the auxiliary equation

$$m^3 + (a - 3)m^2 + (b - a + 2)m + c = 0.$$

Solve

$$11. x^3y''' + x^2y'' - 2xy' + 2y = 0 \quad 12. xy''' + 3y'' = 0$$

13. (Reduction of order) Show that if $y_1(x)$ is a solution of

$$y'' + f(x)y' + g(x)y = 0,$$

then another solution $y_2(x) = u(x)y_1(x)$ can be obtained by solving the following first-order equation in $z = u'$:

$$y_1u'' + (2y_1' + fy_1)u' = 0.$$

14. Show that $y_1 = x$ is a solution of $x^2y'' - x(x+2)y' + (x+2)y = 0$. Using the method in Prob. 13, find another solution. Hint: Divide the equation by x^3 .

15. (Reduction of order) Show that if $y_1(x)$ is a solution of

$$y''' + f(x)y'' + g(x)y' + h(x)y = 0$$

then other solutions $y_2(x) = u(x)y_1(x)$ can be obtained by solving the following second-order equation in $z = u'$:

$$y_1u''' + (3y_1' + fy_1)u'' + (3y_1'' + 2fy_1' + gy_1)u' = 0.$$

Using the method of reduction of order, solve

16. $x^3y''' - 3x^2y'' + 6xy' - 6y = 0, \quad y_1 = x$
17. $x^3y''' - 3xy' + 3y = 0, \quad y_1 = 1/x$
18. $y''' - 3y'' - y' + 3y = 0, \quad y_1 = e^x$
19. $x^3y''' - 3x^2y'' + (6 - x^2)xy' - (6 - x^2)y = 0, \quad y_1 = x$
20. Show that the Wronskian of $y_1 = e^{\lambda_1 x}$, $y_2 = e^{\lambda_2 x}$, $y_3 = e^{\lambda_3 x}$ may be written

$$W = (\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)e^{(\lambda_1 + \lambda_2 + \lambda_3)x}.$$

2.11 Nonhomogeneous Linear Equations

So far we have considered homogeneous linear equations. We shall now discuss linear equations which are nonhomogeneous. For the sake of simplicity we concentrate on second-order equations

$$(1) \quad y'' + f(x)y' + g(x)y = r(x),$$

but it will be obvious that the results in this section can easily be generalized to linear equations of any order.

We shall first show that a general solution $y(x)$ of (1) can be obtained in a simple way from a general solution $y_h(x)$ of the corresponding homogeneous equation

$$(2) \quad y'' + f(x)y' + g(x)y = 0,$$

namely, by adding to y_h any solution \tilde{y} of (1) involving no arbitrary constant:

$$(3) \quad y(x) = y_h(x) + \tilde{y}(x).$$

In fact, since y_h involves two arbitrary constants, so does y and will therefore be a general solution of (1) as defined in Sec. 2.3, provided y is a solution of (1) at all. To show this, we substitute (3) into (1). Then the left-hand side of (1) becomes

$$(y_h + \tilde{y})'' + f(y_h + \tilde{y})' + g(y_h + \tilde{y}).$$

This may be written

$$(y_h'' + fy_h' + gy_h) + \tilde{y}'' + f\tilde{y}' + g\tilde{y}.$$

Since y_h is a solution of (2), the expression in parentheses is zero. Since \tilde{y} satisfies (1), the remaining expression equals $r(x)$, and the statement is proved.

The remaining practical question is how to obtain a solution \tilde{y} of (1). Two corresponding methods will be considered in Secs. 2.12 and 2.16.

We shall now prove that a general solution of an equation (1) with continuous functions f , g , and r represents the totality of all solutions of (1) as follows.

Theorem 1. Suppose that $f(x)$, $g(x)$, and $r(x)$ in (1) are continuous¹¹ functions on an open interval I . Let $Y(x)$ be any solution of (1) on I containing no arbitrary constants. Then $Y(x)$ is obtained from (3) by assigning suitable values to the two arbitrary constants contained in the general solution $y_h(x)$ of (2). In (3), the function $\tilde{y}(x)$ is any solution of (1) on I containing no arbitrary constants.

Proof. We set $Y - \tilde{y} = y^*$. Then

$$y^{*''} + fy^{*''} + gy^* = (Y'' + fY' + gY) - (\tilde{y}'' + f\tilde{y}' + g\tilde{y}) = r - r = 0,$$

that is, y^* is a solution of (2). Now y^* does not contain arbitrary constants. Hence, by Theorem 3, Sec. 2.8, (with Y replaced by y^*), it can be obtained from y_h by assigning suitable values to the arbitrary constants in y_h . From this, since $Y = y^* + \tilde{y}$, the statement follows.

¹¹ This suffices for the existence of y_h on I , as we know, and for the existence of \tilde{y} on I , as will become obvious in Sec. 2.16.

This theorem shows that any solution of (1) containing no arbitrary constants is a particular solution of (1) in the sense of our definition in Sec. 2.3, and we may now formulate our results as follows.

Theorem 2. *A general solution $y(x)$ of the linear nonhomogeneous differential equation (1) is the sum of a general solution $y_h(x)$ of the corresponding homogeneous equation (2) and an arbitrary particular solution $y_p(x)$ of (1):*

$$(4) \quad y(x) = y_h(x) + y_p(x).$$

Problems

1. Show that $y_{p1} = \sin 2x$ and $y_{p2} = \sin 2x - \sin x$ are particular solutions of $y'' + y = -3 \sin 2x$. Find corresponding general solutions (4) and show that one of them can be transformed into the other, so that the choice of y_p is immaterial.
2. Show that two general solutions (4) of (1) which correspond to different y_p can be transformed into each other.
3. Show that if y_1 is a solution of (1) with $r = r_1$ and y_2 is a solution of (1) with $r = r_2$, then $y = y_1 + y_2$ is a solution of (1) with $r = r_1 + r_2$.
4. Show that $y_1 = e^x$ is a solution of $y'' + y = 2e^x$ and $y_2 = x \sin x$ is a solution of $y'' + y = 2 \cos x$. Using Prob. 3, find a general solution of $y'' + y = 2e^x + 2 \cos x$.
5. Show that in Prob. 3, the function $y = ky_1$ is a solution of

$$y'' + f(x)y' + g(x)y = kr_1(x).$$

6. Using Probs. 1 and 5, find a general solution of $y'' + y = \sin 2x$.
7. Extend Theorems 1 and 2 to a third-order differential equation.
8. Extend Theorems 1 and 2 to an equation of any order n .

In each case verify that $y_p(x)$ is a solution of the given differential equation and find a general solution.

9. $y'' + y = 2e^x, \quad y_p = e^x$
10. $y'' - y = 2e^x, \quad y_p = xe^x$
11. $y'' - 3y' + 2y = 2x^2 - 6x + 2, \quad y_p = x^2$
12. $y'' + 4y = -12 \sin 2x, \quad y_p = 3x \cos 2x$
13. $y'' + 9y = 18x, \quad y_p = 2x$
14. $y'' - y = 2 \cos x, \quad y_p = -\cos x$
15. $y'' + y = 2 \sin x, \quad y_p = -x \cos x$

2.12 A Method for Solving Nonhomogeneous Linear Equations

In the last section we have shown that for obtaining a general solution of a linear nonhomogeneous equation we need a particular solution $y_p(x)$ of the equation. How can we find such a function y_p ?

One method is the so-called method of undetermined coefficients. The advantage of this method is that it is much simpler and more economical than another general method to be discussed later (in Sec. 2.16). The disadvantage is that it is applicable only to certain linear equations. But since some of these equations are of great practical importance—for example, in connection with vibrational problems—the method is frequently used in engineering mathematics.

The method is suitable for equations with constant coefficients

$$(1) \quad y'' + ay' + by = r(x)$$

when $r(x)$ is such that the form of a particular solution $y_p(x)$ of (1) may be guessed,¹² for example, r may be a single power of x , a polynomial, an exponential function, a sine or cosine, or a sum of such functions. The method consists in assuming for y_p an expression similar to that of $r(x)$ containing unknown coefficients which are to be determined by inserting y_p and its derivatives into (1).

Let us start with some illustrative examples.

Example 1. Solve the nonhomogeneous equation

$$(2) \quad y'' + 4y = 12.$$

We try $y_p = K$, a constant. Then $y_p'' = 0$. By substitution, $4K = 12$. Thus $y_p = K = 3$. A general solution of the corresponding homogeneous equation is

$$y_h = A \cos 2x + B \sin 2x.$$

Hence, by Theorem 2 in the preceding section, a general solution of (2) is

$$y = y_h + y_p = A \cos 2x + B \sin 2x + 3.$$

Example 2. Solve the nonhomogeneous equation

$$(3) \quad y'' + 4y = 8x^2.$$

We try $y_p = Kx^2$, but this attempt will fail. We have $y_p'' = 2K$. By substitution in (3) we obtain

$$2K + 4Kx^2 = 8x^2.$$

For this to hold for all x , the coefficients of each power of x on both sides (x^2 and x^0) must be the same:

$$4K = 8 \quad \text{and} \quad 2K = 0,$$

which has no solution. Let us try

$$y_p = Kx^2 + Lx + M. \quad \text{Then} \quad y_p'' = 2K.$$

Substitution in (3) yields

$$2K + 4(Kx^2 + Lx + M) = 8x^2.$$

Equating the coefficients of x^2 , x , and x^0 on both sides, we have $4K = 8$, $4L = 0$, $2K + 4M = 0$. Thus $K = 2$, $L = 0$, $M = -1$. Hence $y_p = 2x^2 - 1$, and a general solution of (3) is

$$y = A \cos 2x + B \sin 2x + 2x^2 - 1.$$

¹² The student will see in Sec. 2.16 that our present method actually results from the general method.

Example 3. Solve the nonhomogeneous equation

$$(4) \quad y'' - y' - 2y = 10 \cos x.$$

The student may try $y_p = K \cos x$ and see that this attempt fails. We try

$$y_p = K \cos x + M \sin x.$$

Then

$$y_p' = -K \sin x + M \cos x,$$

$$y_p'' = -K \cos x - M \sin x.$$

By inserting this into (4) we obtain

$$(-3K - M) \cos x + (K - 3M) \sin x = 10 \cos x.$$

By equating the coefficients of $\cos x$ and $\sin x$ on both sides we have

$$-3K - M = 10, \quad K - 3M = 0.$$

Hence $K = -3$, $M = -1$, and $y_p = -3 \cos x - \sin x$. A general solution of the corresponding homogeneous equation is

$$y_h = c_1 e^{-x} + c_2 e^{2x}.$$

Altogether,

$$y = y_h + y_p = c_1 e^{-x} + c_2 e^{2x} - 3 \cos x - \sin x.$$

Table 9. Method of Undetermined Coefficients.

Term in $r(x)$	Choice for y_p
kx^n ($n = 0, 1, \dots$)	$K_n x^n + K_{n-1} x^{n-1} + \dots + K_1 x + K_0$
ke^{px}	Ce^{px}
$k \cos qx$	$K \cos qx + M \sin qx$
$k \sin qx$	iq

In the first column of Table 9 are listed some of the possible terms of $r(x)$. The second column gives the form of y_p , provided the number listed in the last column is not one of the roots of the characteristic equation of the homogeneous equation corresponding to (1).

If $r(x)$ is a sum of functions in the first column, choose for y_p the sum of the functions in the corresponding lines.

Example 4. Find a particular solution of the equation

$$y'' - 3y' + 2y = 4x + e^{3x}.$$

We start from

$$y_p = K_1 x + K_0 + Ce^{3x}.$$

By substitution we find $K_1 = 2$, $K_0 = 3$, $C = 1/2$, and therefore

$$y_p = 2x + 3 + \frac{1}{2}e^{3x}.$$

Modification rule. If the number listed in the last column of Table 9 is a root of the characteristic equation of the homogeneous equation corresponding to (1), the function in the second column of the table must be multiplied by x^m where m is the multiplicity of that root in that equation. (Hence for a second-order equation, m may equal 1 or 2.)

This rule, most likely first discovered by experimenting, can be justified by the general method to be discussed in Sec. 2.16.

Example 5. Solve the nonhomogeneous equation

$$(5) \quad y'' - 2y' + y = e^x + x.$$

The characteristic equation of the corresponding homogeneous equation is

$$(6) \quad \lambda^2 - 2\lambda + 1 = 0.$$

Since 0 is not a root of (6), the function x on the right-hand side of (5) indicates the particular solution choice

$$K_1x + K_0.$$

Since 1 is a double root of (6), that is, a root of multiplicity 2, according to the modification rule the function e^x calls for the particular solution

$$Cx^2e^x \quad (\text{instead of } Ce^x).$$

Altogether,

$$y_p = K_1x + K_0 + Cx^2e^x.$$

Substituting this into (5), we obtain

$$y_p'' - 2y_p' + y_p = 2Ce^x + K_1x - 2K_1 + K_0 = e^x + x.$$

Hence $C = 1/2$, $K_1 = 1$, $K_0 = 2$, and a general solution of (5) is

$$y = y_h + y_p = (c_1x + c_2)e^x + \frac{1}{2}x^2e^x + x + 2.$$

=

Problems

Find a particular solution of the following differential equations.

1. $y'' - 3y' + 2y = 10x$

2. $y'' - y = 1 - 2x$

3. $y'' + y = x^2 + x$

4. $y'' + 2y' + y = 2x^2$

5. $y'' + 5y' + 6y = 9x^4 - x$

6. $y'' - y' - 2y = \sin x$

7. $y'' + y' - 6y = 52 \cos 2x$

8. $y'' + y' - 2y = 3e^x$

9. $y'' + y = 2 \sin x$

10. $y'' - 2y' + 2y = 2e^x \cos x$

11. $y''' + 2y'' - y' - 2y = 1 - 4x^3$

12. $y^{IV} - 5y'' + 4y = 10 \cos x$

Find a general solution of the following differential equations.

13. $y'' + y = -x - x^2$

14. $y'' - y = e^x$

15. $y'' + 4y = e^{-x}$

16. $y'' - y' - 2y = 4 \sin x$

17. $y'' + y = \sin x$

18. $y'' + y' - 2y = e^x$

19. $y'' - 4y' + 3y = e^{3x}$

20. $y'' - 2y' + 2y = e^x \cos x$

The method of undetermined coefficients is useful for certain first-order equations, too, and simpler than the usual method (Sec. 1.7). Using both methods, solve

21. $y' + y = x^4$

22. $y' - 2y = \sin 4x$

23. $y' + 4y = \cos \pi x$

Solve the following initial value problems.

24. $y'' - y = x$, $y(0) = 1$, $y'(0) = 0$

25. $y'' - 2y' + y = 2x^3 - 8x + 4$, $y(0) = 3$, $y'(0) = 3$

26. $y'' - y' - 2y = 10 \sin x$, $y(0) = 1$, $y'(0) = -3$.

27. $y'' - y' - 2y = 3e^{2x}$, $y(0) = 0$, $y'(0) = -2$
 28. $y'' + 2y' + 2y = -2 \cos 2x - 4 \sin 2x$, $y(0) = 1$, $y'(0) = 1$
 29. $y'' + 4y' + 8y = 4 \cos x + 7 \sin x$, $y(0) = 1$, $y'(0) = -1$
 30. $y'' + 2y' + 10y = 4.5 \cos x - \sin x$, $y(0) = 0.5$, $y'(0) = 3$

2.13 Forced Oscillations. Resonance

In Sec. 2.6 we considered free oscillations of a body on a spring as shown in Fig. 65. These motions are governed by the homogeneous equation

$$(1) \quad my'' + cy' + ky = 0$$

where m is the mass of the body, c is the damping constant, and k is the spring modulus. We shall now extend our consideration, assuming that a variable force $r(t)$ acts on the system. In this way we shall become familiar with further interesting facts which are fundamental in engineering mathematics.

We remember that the equation (1) was obtained by considering the forces acting on the body and using Newton's second law. From this it is immediately clear that the differential equation corresponding to the present situation is obtained from (1) by adding the force $r(t)$; this yields

$$my'' + cy' + ky = r(t).$$

Fig. 65. Mass on a spring.

$r(t)$ is called the **input** or **driving force**, and a corresponding solution is called an **output** or a **response of the system to the driving force**. (Cf. also Sec. 1.7.) The resulting motion is called a **forced motion**, in contrast to the **free motion** corresponding to (1), which is a motion in the absence of an external force $r(t)$.

Of particular interest are periodic inputs. We shall consider a sinusoidal input, say,

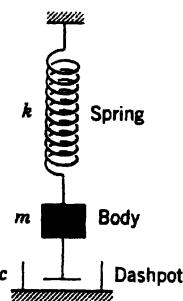
$$r(t) = F_0 \cos \omega t \quad (F_0 > 0, \omega > 0).$$

More complicated periodic inputs will be considered later in Sec. 8.7.

The differential equation now under consideration is

$$(2) \quad my'' + cy' + ky = F_0 \cos \omega t.$$

We want to determine and discuss its general solution, which represents the general output of our vibrating system. Since a general solution of the corresponding homogeneous equation (1) is known from Sec. 2.6, we must now determine a particular solution $y_p(t)$ of (2).



This can be done by the method of undetermined coefficients (Sec. 2.12), starting from

$$(3) \quad y_p(t) = a \cos \omega t + b \sin \omega t.$$

By differentiating this function we have

$$\dot{y}_p = -\omega a \sin \omega t + \omega b \cos \omega t, \quad \ddot{y}_p = -\omega^2 a \cos \omega t - \omega^2 b \sin \omega t.$$

By substituting these expressions into (2) and collecting the cosine and the sine terms we immediately obtain

$$[(k - m\omega^2)a + \omega c b] \cos \omega t + [-\omega c a + (k - m\omega^2)b] \sin \omega t = F_0 \cos \omega t.$$

By equating the coefficients of the cosine and sine terms on both sides we have

$$(4) \quad \begin{aligned} (k - m\omega^2)a + \omega c b &= F_0 \\ -\omega c a + (k - m\omega^2)b &= 0. \end{aligned}$$

The solution of this system of equations in the unknowns a and b is (cf. Sec. 0.3)

$$a = F_0 \frac{k - m\omega^2}{(k - m\omega^2)^2 + \omega^2 c^2}, \quad b = F_0 \cdot \frac{\omega c}{(k - m\omega^2)^2 + \omega^2 c^2}$$

provided the denominator is not zero. If we set $\sqrt{k/m} = \omega_0 (> 0)$ as in Sec. 2.6, this becomes

$$(5) \quad a = F_0 \frac{m(\omega_0^2 - \omega^2)}{m^2(\omega_0^2 - \omega^2)^2 + \omega^2 c^2}, \quad b = F_0 \frac{\omega c}{m^2(\omega_0^2 - \omega^2)^2 + \omega^2 c^2}.$$

We thus obtain the general solution

$$(6) \quad y(t) = y_h(t) + y_p(t),$$

where y_h is a general solution of (1) and y_p is given by (3) with coefficients (5).

We shall now discuss the behavior of the mechanical system, distinguishing between the two cases $c = 0$ (no damping) and $c > 0$, which will correspond to two different types of the output.

Case 1. Undamped forced oscillations. If there is no damping, then $c = 0$, and assuming that $\omega \neq \omega_0$, we obtain from (3) and (5)

$$(7) \quad y_p(t) = \frac{F_0}{m(\omega_0^2 - \omega^2)} \cos \omega t = \frac{F_0}{k[1 - (\omega/\omega_0)^2]} \cos \omega t.$$

From this and (6*) in Sec. 2.6 we have the general solution

$$(8) \quad y(t) = C \cos(\omega_0 t - \delta) + \frac{F_0}{m(\omega_0^2 - \omega^2)} \cos \omega t.$$

This output represents a superposition of two harmonic oscillations; the frequencies are the "natural frequency" $\omega_0/2\pi$ [cycles/sec] of the system (that

is, the frequency of the free undamped motion) and the frequency $\omega/2\pi$ of the input.

From (7) we see that the maximum amplitude of y_p is

$$(9) \quad a_0 = \frac{F_0}{k} \rho \quad \text{where} \quad \rho = \frac{1}{1 - (\omega/\omega_0)^2}.$$

It depends on ω and ω_0 . As $\omega \rightarrow \omega_0$, the quantities ρ and a_0 tend to infinity. This phenomenon of excitation of large oscillations by matching input and natural frequencies ($\omega = \omega_0$) is known as **resonance**, and is of basic importance in the study of vibrating systems (cf. below). The quantity ρ is called the **resonance factor** (Fig. 66). From (9) we see that ρ/k is the ratio of the amplitudes of the function y_p and the input.

In the case of resonance the equation (2) becomes

$$(10) \quad \ddot{y} + \omega_0^2 y = \frac{F_0}{m} \cos \omega_0 t.$$

From the modification rule in the preceding section we conclude that a particular solution of (10) has the form

$$y_p(t) = t(a \cos \omega_0 t + b \sin \omega_0 t).$$

By substituting this into (10) we find $a = 0$, $b = F_0/2m\omega_0$, and (Fig. 67)

$$(11) \quad y_p(t) = \frac{F_0}{2m\omega_0} t \sin \omega_0 t.$$

Another interesting and highly important type of oscillation is obtained when ω is close to ω_0 . Take, for example, the particular solution [cf. (8)]

$$(12) \quad y(t) = \frac{F_0}{m(\omega_0^2 - \omega^2)} (\cos \omega t - \cos \omega_0 t) \quad (\omega \neq \omega_0)$$

corresponding to the initial conditions $y(0) = 0$, $\dot{y}(0) = 0$. From (11a) in Sec. 0.1 it follows that

$$y(t) = \frac{2F_0}{m(\omega_0^2 - \omega^2)} \sin \frac{\omega_0 + \omega}{2} t \sin \frac{\omega_0 - \omega}{2} t.$$

Since ω is close to ω_0 the difference $\omega_0 - \omega$ is small, so that the period of the last sine function is large, and we obtain an oscillation of the type shown in Fig. 68.

Case 2. Damped forced oscillations. If there is damping, then $c > 0$, and we know from Sec. 2.6 that the general solution y_h of (1) approaches zero as t approaches infinity (practically, after a sufficiently long time); that is, the general solution (6) of (2) now represents the **transient solution** and tends to the **steady-state solution** y_p . Hence, after a sufficiently long time, the output corresponding to a purely sinusoidal input will approximate a harmonic

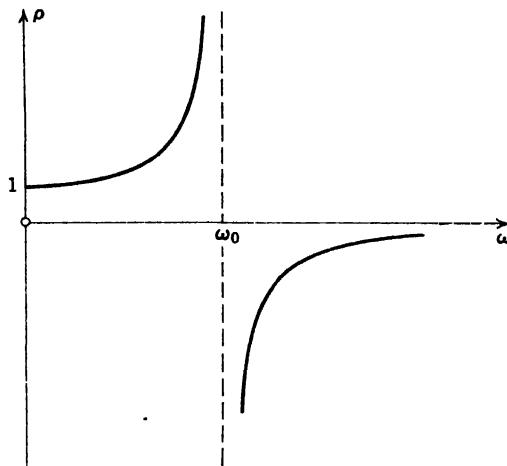
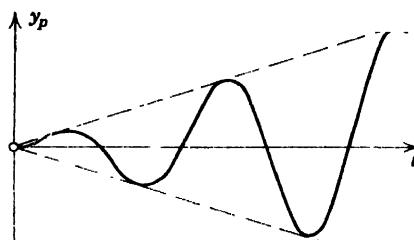
Fig. 66. Resonance factor $\rho(\omega)$.

Fig. 67. Particular solution in the case of resonance.

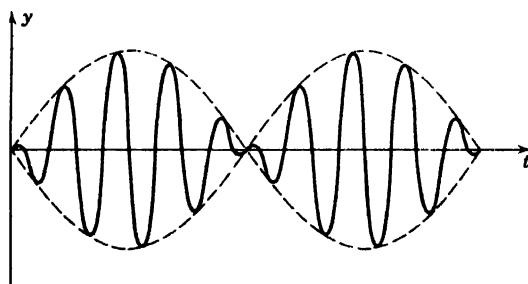


Fig. 68. Forced undamped oscillation when the difference of the input and natural frequencies is small ("beats").

oscillation whose frequency is that of the input. This characterizes the practical situation because no physical system is completely undamped.

While in the undamped case the amplitude y_p approaches infinity as ω approaches ω_0 , this will not happen in the present case. The amplitude will always be finite, but may have a maximum for some ω , depending on c . This may be called *practical resonance*. It is of great importance because it shows that some input may excite oscillations with such a large amplitude that the system may be destroyed. Such cases happened in practice, in particular in earlier times when less was known about resonance. Cars, ships, airplanes, and bridges are vibrating mechanical systems, and it is sometimes rather difficult to find constructions which are completely free of undesired resonance effects.

To investigate the amplitude of y_p as a function of ω , we write (3) in the form

$$(13) \quad y_p(t) = C^* \cos(\omega t - \eta)$$

where, according to (5),

$$(14) \quad \begin{aligned} C^*(\omega) &= \sqrt{a^2 + b^2} = \frac{\omega c}{\sqrt{m^2(\omega_0^2 - \omega^2)^2 + \omega^2 c^2}}, \\ \tan \eta &= \frac{c}{a} = \frac{\omega c}{m(\omega_0^2 - \omega^2)} \end{aligned}$$

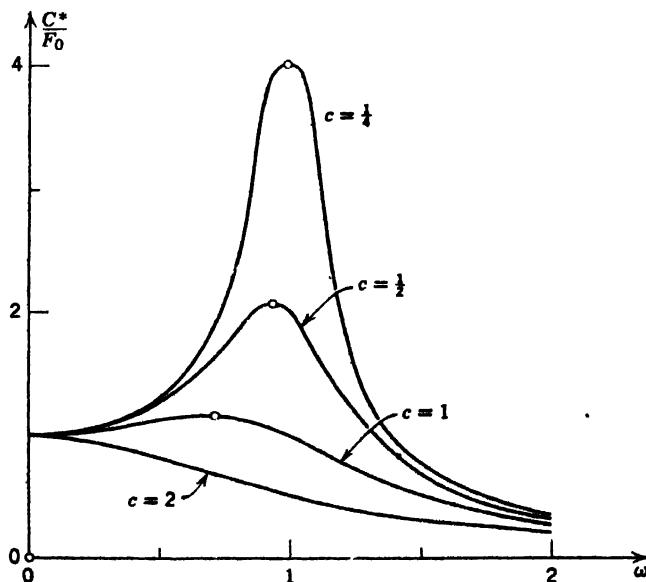


Fig. 69. Amplification C^*/F_0 as a function of ω for $m = 1$, $k = 1$, and various values of the damping constant c .

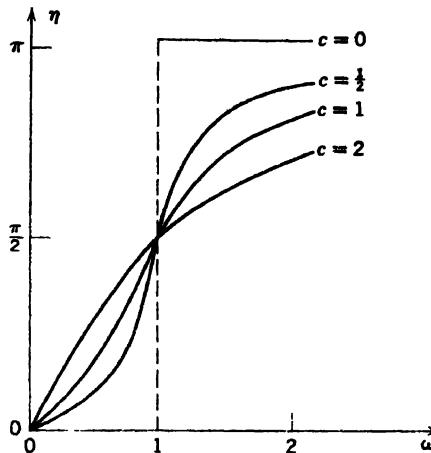


Fig. 70. Phase lag η as a function of ω for $m = 1$, $k = 1$, and various values of the damping constant c .

Let us determine the maximum of $C^*(\omega)$. By equating $dC^*/d\omega$ to zero we find

$$[-2m^2(\omega_0^2 - \omega^2) + c^2]\omega = 0.$$

The expression in brackets is zero for

$$(15) \quad c^2 = 2m^2(\omega_0^2 - \omega^2).$$

For sufficiently large damping ($c^2 > 2m^2\omega_0^2 = 2mk$) the equation (15) has no real solution, and C^* decreases in a monotone way as ω increases (Fig. 69). If $c^2 \leq 2mk$, the equation (15) has a real solution $\omega = \omega_{\max}$, which increases as c decreases and approaches ω_0 as c approaches zero. The amplitude $C^*(\omega)$ has a maximum at $\omega = \omega_{\max}$, and by inserting $\omega = \omega_{\max}$ into (14) we find that

$$(16) \quad C^*(\omega_{\max}) = \frac{2mF_0}{c\sqrt{4m^2\omega_0^2 - c^2}}.$$

We see that $C^*(\omega_{\max})$ is finite when $c > 0$. Since $dC^*(\omega_{\max})/dc < 0$ when $c^2 < 2mk$, the value of $C^*(\omega_{\max})$ increases as c ($\leq \sqrt{2mk}$) decreases and approaches infinity as c approaches zero, in agreement with our result in Case 1. Figure 69 shows the amplification C^*/F_0 (ratio of the amplitudes of output and input) as a function of ω for $m = 1$, $k = 1$, and various values of the damping constant c .

The angle η in (14) is called the phase angle or phase lag (Fig. 70) because it measures the lag of the output with respect to the input. If $\omega < \omega_0$, then $\eta < \pi/2$; if $\omega = \omega_0$, then $\eta = \pi/2$, and if $\omega > \omega_0$, then $\eta > \pi/2$.

Problems

Find and graph the oscillations of the mechanical systems corresponding to the following differential equations.

1. $\ddot{y} + y = 3 \cos 2t$
 3. $\ddot{y} + 16y = \cos t - \sin t$

2. $\ddot{y} + 4y = \sin t$
 4. $\ddot{y} + 9y = 6 \cos 3t$

Solve the following initial value problems.

5. $\ddot{y} + 25y = 24 \sin t, \quad y(0) = 1, \quad \dot{y}(0) = 1$
 6. $\ddot{y} + y = \cos \omega t, \quad y(0) = y_0, \quad \dot{y}(0) = v_0, \quad \omega^2 \neq 1$
 7. $\ddot{y} + \omega_0^2 y = \cos \omega t, \quad y(0) = y_0, \quad \dot{y}(0) = v_0, \quad \omega^2 \neq \omega_0^2$
 8. For what initial conditions does the solution of Prob. 6 represent an oscillation whose frequency equals the frequency of the input?
 9. Can we find initial conditions such that the solution of Prob. 7 represents an oscillation whose frequency equals the natural frequency of the system?
 10. Solve the initial value problem $\ddot{y} + y = \cos \omega t, \quad \omega^2 \neq 1, \quad y(0) = 0, \quad \dot{y}(0) = 0$, and graph the maximum amplitude as a function of ω .
 11. Show that the solution in Prob. 10 may be written

$$y(t) = 2(1 - \omega^2)^{-1} \sin [0.5(1 + \omega)t] \sin [0.5(1 - \omega)t],$$

and plot good graphs of $y(t)$ with $\omega = 0.5, 0.9, 1.1, 2$.

12. In (12) let ω approach ω_0 . Show that this leads to a solution of the form (11).

Find the steady-state oscillations of the mechanical systems corresponding to the following differential equations.

13. $\ddot{y} + 3\dot{y} + 2y = 10 \sin t$

14. $\ddot{y} + 4\dot{y} + 3y = 65 \cos 2t$

15. $\ddot{y} + 2\dot{y} + 2y = 85 \cos 3t$

16. $\ddot{y} + 4\dot{y} + 5y = 37.7 \sin 4t$

17. Find a particular solution of the differential equation

$$m\ddot{y} + c\dot{y} + ky = b \sin \omega t \quad (c \neq 0).$$

Find the transient oscillations of the mechanical systems governed by the following equations.

18. $\ddot{y} + 2\dot{y} + 5y = 10 \sin t$

19. $2\ddot{y} + \dot{y} + 4y = 10 \cos t$

20. $\ddot{y} + 2\dot{y} + 2y = 5 \cos t$

21. $\ddot{y} + 2\dot{y} + y = 100 \sin 3t$

Solve the following initial value problems.

22. $\ddot{y} + 10\dot{y} + 29y = 10 \cos t + 28 \sin t, \quad y(0) = 1, \quad \dot{y}(0) = -4$

23. $0.5\ddot{y} + 2\dot{y} + 2.5y = \cos 5t + \sin 5t, \quad y(0) = -0.1, \quad \dot{y}(0) = 0$

24. $\ddot{y} + y = -9 \sin 2t, \quad y(0) = 1, \quad \dot{y}(0) = 0$

25. $\ddot{y} + 2\dot{y} + 2y = \sin 2t - 2 \cos 2t, \quad y(0) = 0, \quad \dot{y}(0) = 0$

2.14 Electric Circuits

The last section was devoted to the study of a mechanical system which is of great practical interest. We shall now consider a similarly important

electrical system, which may be regarded as a basic building block in electrical networks. This consideration will also provide a striking example of the important fact that entirely different physical systems may correspond to the same differential equation, and thus illustrate the role which mathematics plays in unifying various phenomena of entirely different physical nature. We shall obtain a correspondence between mechanical and electrical systems which is not merely qualitative but strictly quantitative in the sense that to a given mechanical system we can construct an electric circuit whose current will give the exact values of the displacement in the mechanical system when suitable scale factors are introduced. The practical importance of such an analogy between mechanical and electrical systems is almost obvious. The analogy may be used for constructing an "electrical model" of a given mechanical system; in many cases this will be an essential simplification, because electric circuits are easy to assemble and currents and voltages are easy to measure, whereas the construction of a mechanical model may be complicated and expensive, and the measurement of displacements will be time-consuming and inaccurate.

We shall consider the "RLC-circuit" in Fig. 71. Simpler circuits were considered in Sec. 1.9, and it will be assumed that the student is familiar with those simpler cases, because they will immediately provide the starting point for obtaining the differential equation of our present circuit. We remember that the differential equations of those circuits were obtained by equating the sum of the voltage drops across the elements of the circuits to the impressed electromotive force. This was done by using Kirchhoff's second law, which therefore plays a role in circuits similar to that of Newton's second law (Sec. 2.6) in mechanical systems.

To obtain the differential equation of the circuit in Fig. 71, we may use the equation of an *RC*-circuit [cf. (7), Sec. 1.9] and add to it the voltage drop $L\dot{I}$ across the inductor in our present circuit. This yields

$$(1') \quad L\ddot{I} + RI + \frac{1}{C} \int I dt = E(t) = E_0 \sin \omega t.$$

Differentiating this equation with respect to t , we obtain the differential equation of our *RLC*-circuit in the form

$$(1) \quad L\ddot{I} + R\dot{I} + \frac{1}{C} I = E_0 \omega \cos \omega t.$$

This equation and the equation (2) in Sec. 2.13 are essentially the same. This shows that our RLC-circuit is the electrical analogue of the mechanical system in Sec. 2.13, and we note the following correspondence between the

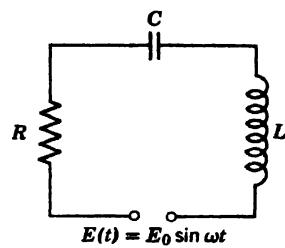


Fig. 71. *RLC*-circuit.

electrical and mechanical quantities:

$$\text{Inductance } L \leftrightarrow \text{mass } m$$

$$\text{Resistance } R \leftrightarrow \text{damping constant } c$$

$$\text{Reciprocal of capacitance } 1/C \leftrightarrow \text{spring modulus } k$$

$$\left. \begin{array}{l} \text{Derivative of} \\ \text{electromotive force } E_0 \omega \cos \omega t \end{array} \right\} \leftrightarrow \text{driving force } F_0 \cos \omega t \text{ (input)}$$

$$\text{Current } I(t) \leftrightarrow \text{displacement } y(t) \text{ (output)}$$

To obtain a particular solution of (1) we may proceed as in Sec. 2.13. By substituting

$$(2) \quad I_p(t) = a \cos \omega t + b \sin \omega t$$

into (1) we obtain

$$(3) \quad a = \frac{E_0 S}{R^2 + S^2}, \quad b = \frac{E_0 R}{R^2 + S^2}$$

where S is the so-called **reactance**, given by the expression

$$(4) \quad S = \omega L - \frac{1}{\omega C}.$$

In any practical case, $R \neq 0$ so that the denominator in (3) is not zero. The result is that (2), with a and b given by (3), is a particular solution of (1).

Using (3), we may write I_p in the form

$$(5) \quad I_p(t) = I_0 \sin(\omega t - \theta)$$

where (cf. Sec. 0.1)

$$I_0 = \sqrt{a^2 + b^2} = \frac{E_0}{\sqrt{R^2 + S^2}}, \quad \tan \theta = -\frac{a}{b} = \frac{S}{R}.$$

The quantity $\sqrt{R^2 + S^2}$ is called the **impedance**.

The characteristic equation

$$\lambda^2 + \frac{R}{L} \lambda + \frac{1}{LC} = 0$$

of the homogeneous differential equation corresponding to (1) has the roots $\lambda_1 = -\alpha + \beta$ and $\lambda_2 = -\alpha - \beta$ where

$$\alpha = \frac{R}{2L}, \quad \beta = \frac{1}{2L} \sqrt{R^2 - \frac{4}{C}}.$$

As in Sec. 2.13 we conclude that if $R > 0$ (which, of course, is true in any practical case) the general solution $I_h(t)$ of the homogeneous equation approaches zero as t approaches infinity (practically: after a sufficiently long

time). Hence, the transient current $I = I_h + I_s$, tends to the steady state current I_s , and *after a sufficiently long time the output will practically be a harmonic oscillation, which is given by (5) and whose frequency is that of the input.*

Problems

Find the steady-state current in the *RLC*-circuit in Fig. 72 where

1. $R = 20$ ohms, $L = 10$ henrys, $C = 0.05$ farad, $E = 50 \sin t$ volts
2. $R = 240$ ohms, $L = 40$ henrys, $C = 10^{-3}$ farad, $E = 369 \sin 10t$ volts
3. $R = 40$ ohms, $L = 10$ henrys, $C = 0.02$ farad, $E = 800 \cos 5t$ volts

Find the transient current in the *RLC*-circuit in Fig. 72 where

4. $R = 200$ ohms, $L = 100$ henrys, $C = 5 \cdot 10^{-3}$ farad, $E = 2500 \sin t$ volts
5. $R = 20$ ohms, $L = 5$ henrys, $C = 0.01$ farad, $E = 850 \sin 4t$ volts
6. $R = 16$ ohms, $L = 8$ henrys, $C = 0.125$ farad, $E = 300 \cos 2t$ volts
7. Show that if at $t = 0$ the charge in the capacitor in Fig. 72 is $Q(0)$, then

$$i(0) = \frac{E(0)}{L} - \frac{R}{L} I(0) - \frac{Q(0)}{LC} \quad (LC \neq 0).$$

Find the current $I(t)$ in the *LC*-circuit in Fig. 73, assuming zero initial current and charge, and

8. $L = 10$ henrys, $C = 4 \cdot 10^{-3}$ farad, $E = 250$ volts
9. $L = 1$ henry, $C = 0.25$ farad, $E = 90 \cos t$ volts
10. $L = 2$ henrys, $C = 5 \cdot 10^{-3}$ farad, $E = 210 \sin 4t$ volts

Find the current $I(t)$ in the *RLC*-circuit in Fig. 72, assuming zero initial current and charge, and

11. $R = 80$ ohms, $L = 20$ henrys, $C = 0.01$ farad, $E = 100$ volts
12. $R = 160$ ohms, $L = 20$ henrys, $C = 2 \cdot 10^{-3}$ farad, $E = 481 \sin 10t$ volts
13. $R = 6$ ohms, $L = 1$ henry, $C = 0.04$ farad, $E = 24 \cos 5t$ volts
14. Show that if $E(t)$ in Fig. 73 has a jump of magnitude J at $t = a$, then $i(t)$ has a jump of magnitude J/L at $t = a$ while $I(t)$ is continuous at $t = a$.

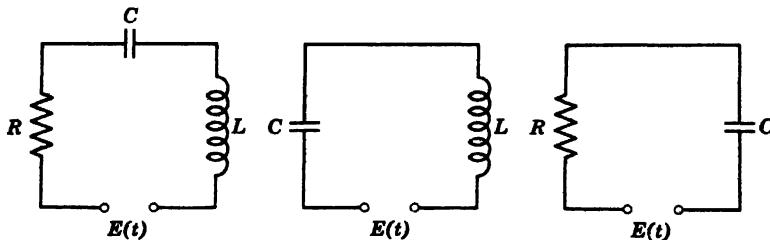


Fig. 72. *RLC*-circuit.

Fig. 73. *LC*-circuit.

Fig. 74. *RC*-circuit.

Using the result of Prob. 14, find the current $I(t)$ in the LC -circuit in Fig. 73, assuming $L = 1$ henry, $C = 1$ farad, zero initial current and charge, and

15. $E = t$ when $0 < t < a$ and $E = 0$ when $t > a$
16. $E = 1$ when $0 < t < a$ and $E = 0$ when $t > a$
17. $E = 1 - e^{-t}$ when $0 < t < \pi$ and $E = 0$ when $t > \pi$
18. $E = t$ when $0 < t < a$ and $E = a$ when $t > a$
19. Show that if the initial charge in the capacitor in Fig. 74 is $q(0)$, the initial current in the RC -circuit is

$$I(0) = \frac{E(0)}{R} - \frac{q(0)}{RC}.$$

20. Show that if $E(t)$ in Fig. 74 has a jump of magnitude J at $t = a$, then the current $I(t)$ in the circuit has a jump of magnitude J/R at $t = a$.

Using the result of Prob. 20, find the current $I(t)$ in the RC -circuit in Fig. 74, assuming $R = 1$ ohm, $C = 1$ farad, zero initial charge on the capacitor, and

21. $E = t$ when $0 < t < a$ and $E = a$ when $t > a$
22. $E = t$ when $0 < t < a$ and $E = 0$ when $t > a$
23. $E = t + 1$ when $0 < t < a$ and $E = 0$ when $t > a$
24. $E = t + 1$ when $0 < t < a$ and $E = a + 1$ when $t > a$

2.15 Complex Method for Obtaining Particular Solutions

Given an equation of the form (1) in Sec. 2.14, for example,

$$(1) \quad \ddot{I} + \dot{I} + 2I = 6 \cos t,$$

we know that we can obtain a particular solution $I_p(t)$ by the method of undetermined coefficients, that is, by substituting

$$I_p(t) = a \cos t + b \sin t$$

into (1) and determining a and b . The result will be

$$I_p(t) = 3 \cos t + 3 \sin t.$$

Engineers often prefer a simple and elegant complex method for obtaining $I_p(t)$. This method proceeds as follows.

We start from the differential equation

$$(2) \quad I + \dot{I} + 2I = 6e^{it} \quad (i = \sqrt{-1}).$$

This equation suggests to determine a complex particular solution of the form

$$(3) \quad I_p^*(t) = Ke^{it}.$$

By substituting this function and its derivatives

$$I_p^* = iKe^{it}, \quad \dot{I}_p^* = -Ke^{it}$$

into the equation (2) we have

$$(-1 + i + 2)Ke^{it} = 6e^{it}.$$

Solving for K , we obtain [cf. (4) in Sec. 0.4]

$$\frac{6}{1+i} = \frac{6(1-i)}{(1+i)(1-i)} = 3i.$$

Hence

$$I_p^*(t) = (3 - 3i)e^{it} = (3 - 3i)(\cos t + i \sin t)$$

is a solution of (2); here the last expression is obtained by means of the Euler formula (Sec. 2.4). The real part of I_p^* is

$$I_p(t) = 3 \cos t + 3 \sin t$$

and this function is a solution of the real part of the differential equation (2), that is, of the given differential equation (1). In fact, the function is identical with that obtained above. This illustrates the practical procedure of the complex method.

Our equation (1) is a particular case of the equation [cf. (1), Sec. 2.14]

$$(4) \quad L\ddot{I} + R\dot{I} + \frac{1}{C}I = E_0\omega \cos \omega t$$

which we shall now consider, assuming that $R \neq 0$. The corresponding complex equation is

$$(5) \quad L\ddot{I} + R\dot{I} + \frac{1}{C}I = E_0\omega e^{i\omega t}.$$

The function on the right suggests a particular solution of the form

$$(6) \quad I_p^*(t) = Ke^{i\omega t} \quad (i = \sqrt{-1}).$$

By substituting this function and its derivatives

$$I_p^* = i\omega K e^{i\omega t} \quad \dot{I}_p^* = -\omega^2 K e^{i\omega t}$$

into (5) we immediately have

$$\left(-\omega^2 L + i\omega R + \frac{1}{C} \right) K e^{i\omega t} = E_0 \omega e^{i\omega t}.$$

Solving this equation for K , we obtain

$$(7) \quad K = \frac{E_0 \omega}{\frac{1}{C} - \omega^2 L + i\omega R} = \frac{E_0}{iZ}$$

where Z is the so-called complex impedance, given by

$$(8) \quad Z = R + i\left(\omega L - \frac{1}{\omega C}\right).$$

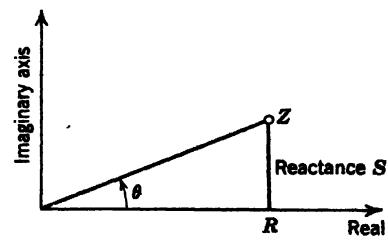


Fig. 75. Complex impedance Z .

It follows that (6) with K given by (7) is a solution of (5).

We see that the imaginary part of Z is the reactance S defined by (4), Sec. 2.14, and $|Z|$ is the impedance defined in connection with (5), Sec. 2.14. Thus

$$Z = |Z| e^{i\theta} \quad \text{where} \quad \tan \theta = \frac{S}{R},$$

and (6) with K according to (7) can be written in the form

$$I_p^*(t) = \frac{E_0}{iZ} e^{i\omega t} = -i \frac{E_0}{|Z|} e^{i(\omega t - \theta)}.$$

The real part is

$$(9) \quad I_p(t) = \frac{E_0}{|Z|} \sin(\omega t - \theta) = \frac{E_0}{\sqrt{R^2 + S^2}} \sin(\omega t - \theta),$$

and this is a solution of the real part of the differential equation (5), that is, $I_p(t)$ is a solution of (4). We see that this function is identical with (5) in the previous section. Since

$$\sin(\omega t - \theta) = \sin \omega t \cos \theta - \cos \omega t \sin \theta$$

and furthermore

$$\cos \theta = \frac{\operatorname{Re} Z}{|Z|} = \frac{R}{\sqrt{R^2 + S^2}}, \quad \sin \theta = \frac{\operatorname{Im} Z}{|Z|} = \frac{S}{\sqrt{R^2 + S^2}},$$

the solution can be written in the form

$$(10) \quad I_p(t) = \frac{E_0}{R^2 + S^2} [R \sin \omega t - S \cos \omega t].$$

This is identical with (2) in the previous section, where a and b are given by (3) in that section.

Problems

Using the complex method, find the steady-state current $I_p(t)$ in the circuit governed by (4), where

1. $R = 20, L = 10, C = 0.05, E_0 = 5, \omega = 2$
2. $R = 100, L = 100, C = 10^{-3}, E_0 = 3, \omega = 4$
3. $R = 50, L = 25, C = 0.01, E_0 = 500, \omega = 3$
4. $R = 10, L = 20, C = 0.1, E_0 = 4, \omega = 2$
5. $R = 25, L = 15, C = 0.05, E_0 = 100, \omega = 4$

6. Find the complex impedance Z and the reactance S in Prob. 5.

Using the complex method, find the steady-state output of the following equations.

- | | |
|---|--|
| 7. $\ddot{y} + 3\dot{y} + 16y = 24 \cos 4t$ | 8. $\ddot{y} + \dot{y} + 4y = 8 \sin 2t$ |
| 9. $\ddot{y} + 5\dot{y} + 8y = 25 \cos 10t$ | 10. $\ddot{y} + \dot{y} + 9y = -3 \sin 3t$ |

2.16 General Method for Solving Nonhomogeneous Equations

The method of undetermined coefficients in Sec. 2.12 is a simple procedure for determining particular solutions of nonhomogeneous linear equations, and we have seen in Secs. 2.13–2.15 that it has important applications in connection with vibrations. However, it is limited to certain simpler types of differential equations. The method to be discussed in this section will be more general but more complicated than that method.

We consider equations of the form

$$(1) \quad y'' + f(x)y' + g(x)y = r(x)$$

assuming that f , g , and r are continuous on an open interval I . We shall obtain a particular solution of (1) by using the method of **variation of parameters**¹³ as follows.

We know that the corresponding homogeneous equation

$$(2) \quad y'' + f(x)y' + g(x)y = 0$$

has a general solution $y_h(x)$ on I , which is of the form

$$y_h(x) = c_1 y_1(x) + c_2 y_2(x).$$

The method consists in replacing c_1 and c_2 by functions $u(x)$ and $v(x)$ to be determined such that the resulting function

$$(3) \quad y_p(x) = u(x)y_1(x) + v(x)y_2(x)$$

is a particular solution of (1) on I . By differentiating (3) we obtain

$$y_p' = u'y_1 + uy_1' + v'y_2 + vy_2'.$$

We shall see that we can determine u and v such that

$$(4) \quad u'y_1 + v'y_2 = 0.$$

This reduces the expression for y_p' to the form

$$(5) \quad y_p' = uy_1' + vy_2'.$$

By differentiating this function we obtain

$$(6) \quad y_p'' = u'y_1' + uy_1'' + v'y_2' + vy_2''.$$

By substituting (3), (5), and (6) into (1) and collecting terms containing u and terms containing v we readily obtain

$$u(y_1'' + fy_1' + gy_1) + v(y_2'' + fy_2' + gy_2) + u'y_1' + v'y_2' = r.$$

¹³ The student may wish to review the simpler application of the method in Sec. 1.8.

Since y_1 and y_2 are solutions of the homogeneous equation (2), this reduces to

$$u'y_1' + v'y_2' = r.$$

Equation (4) is

$$u'y_1 + v'y_2 = 0.$$

This is a system of two linear algebraic equations for the unknown functions u' and v' . The solution is obtained by Cramer's rule (cf. Sec. 0.3). We find

$$(7) \quad u' = -\frac{y_2r}{W}, \quad v' = \frac{y_1r}{W},$$

where

$$W = y_1y_2' - y_1'y_2$$

is the Wronskian of y_1 and y_2 . Clearly, $W \neq 0$, because the functions y_1 and y_2 constitute a fundamental system. By integrating (7) we have

$$u = -\int \frac{y_2r}{W} dx, \quad v = \int \frac{y_1r}{W} dx.$$

These integrals exist because $r(x)$ is continuous. By substituting these expressions for u and v into (3) we obtain the desired solution of (1),

$$(8) \quad y_p(x) = -y_1 \int \frac{y_2r}{W} dx + y_2 \int \frac{y_1r}{W} dx.$$

Note that if the constants of integration in (8) are left arbitrary, then (8) represents a general solution of (1).

Example. Solve the differential equation

$$(9) \quad y'' + y = \sec x.$$

The method of undetermined coefficients cannot be applied. The functions

$$y_1 = \cos x, \quad y_2 = \sin x$$

constitute a fundamental system of the homogeneous equation, and their Wronskian is

$$W(y_1, y_2) = \cos x \cos x - (-\sin x) \sin x = 1.$$

From (8) we thus obtain the following solution of (9),

$$\begin{aligned} y_p &= -\cos x \int \sin x \sec x dx + \sin x \int \cos x \sec x dx \\ &= \cos x \ln(\cos x) + x \sin x. \end{aligned}$$

Hence the general solution of the differential equation (9) is

$$y = y_h + y_p = [c_1 + \ln(\cos x)] \cos x + (c_2 + x) \sin x.$$

Problems

Find a general solution of the following equations.

1. $y'' + y = \csc x$

3. $y'' - 4y' + 4y = e^{2x}/x$

2. $y'' + 9y = \sec 3x$

4. $y'' + 6y' + 9y = e^{-3x}/(x^2 + 1)$

- | | |
|---------------------------------------|--|
| 5. $y'' + 2y' + y = e^{-x} \cos x$ | 6. $y'' + 2y' + y = e^{-x} \ln x$ |
| 7. $y'' - 2y' + y = x^3 e^x$ | 8. $y'' - 2y' + y = e^x/x^3$ |
| 9. $y'' + 2y' + 2y = e^{-x}/\cos^3 x$ | 10. $y'' - 4y' + 5y = e^{2x}/\sin x$ |
| 11. $x^2 y'' + xy' - y = 4$ | 12. $x^2 y'' - 2y = 3x^2$ |
| 13. $x^2 y'' - 2xy' + 2y = x^4$ | 14. $x^2 y'' + xy' - 0.25y = 1/x$ |
| 15. $x^2 y'' - 2xy' + 2y = 1/x^2$ | 16. $x^2 y'' - 2xy' + 2y = x^3 \cos x$ |
| 17. $x^2 y'' - 4xy' + 6y = 1/x^4$ | 18. $x^2 y'' - 4xy' + 6y = x^4 \sin x$ |
| 19. $xy'' - y' = 2x^2 e^x$ | 20. $x^2 y'' + xy' - y = x^3 e^x$ |

2.17 Numerical Methods for Second-Order Differential Equations

We shall now consider two important numerical methods for solving initial value problems of the form

$$(1) \quad y'' = f(x, y, y'), \quad y(x_0) = y_0, \quad y'(x_0) = y_0'$$

where f is assumed to be such that the problem has a unique solution on some interval containing x_0 .

We shall obtain approximate values of the solution $y(x)$ of (1) at equidistant points $x_1 = x_0 + h, x_2 = x_0 + 2h, \dots$; these values will be denoted by y_1, y_2, \dots , respectively. Similarly, approximate values of the derivative $y'(x)$ at those points will be denoted by y'_1, y'_2, \dots , respectively.

In Sec. 1.13 we have seen that numerical methods for solving first-order differential equations can be obtained by the use of the Taylor series

$$(2) \quad y(x+h) = y(x) + hy'(x) + \frac{h^2}{2} y''(x) + \frac{h^3}{3!} y'''(x) + \dots$$

In the present consideration we shall use this series for the same purpose, and we shall also need the Taylor series for the derivative,

$$(3) \quad y'(x+h) = y'(x) + hy''(x) + \frac{h^2}{2} y'''(x) + \dots$$

The roughest numerical method is obtained by neglecting the terms containing y'' and the further terms in (2) and (3); this yields the approximations

$$y(x+h) \approx y(x) + hy'(x) + \frac{h^2}{2} y''(x),$$

$$y'(x+h) \approx y'(x) + hy''(x).$$

In the first step of the method we compute

$$y_0'' = f(x_0, y_0, y_0')$$

from (1), then

$$y_1 = y_0 + hy_0' + \frac{h^2}{2} y_0''$$

which approximates $y(x_1) = y(x_0 + h)$, and furthermore

$$y_1' = y_0' + hy_0''$$

which will be needed in the next step. In the second step we compute

$$y_1'' = f(x_1, y_1, y_1')$$

from (1), then

$$y_2 = y_1 + hy_1' + \frac{h^2}{2} y_1''$$

which approximates $y(x_2) = y(x_0 + 2h)$, and furthermore

$$y_2' = y_1' + hy_1''.$$

In the $(n + 1)$ th step we compute

$$y_n'' = f(x_n, y_n, y_n')$$

from (1), then the new value

$$(4a) \quad y_{n+1} = y_n + hy_n' + \frac{h^2}{2} y_n''$$

which is an approximation for $y(x_{n+1})$, and furthermore

$$(4b) \quad y_{n+1}' = y_n' + hy_n''$$

which is an approximation for $y'(x_{n+1})$ needed in the next step.

Note that, geometrically speaking, this method is an approximation of the curve of $y(x)$ by portions of parabolas.

Example 1. Apply (4) to the initial value problem

$$(5) \quad y'' = \frac{1}{2}(x + y + y' + 2), \quad y(0) = 0, \quad y'(0) = 0,$$

choosing $h = 0.2$. Here (4) becomes

$$\begin{aligned} y_{n+1} &= y_n + 0.2y_n' + 0.02y_n'' \\ y_{n+1}' &= y_n' + 0.2y_n'' \end{aligned} \quad \text{where} \quad y_n = \frac{1}{2}(x_n + y_n + y_n' + 2).$$

The computations are as follows:

n	x_n	y_n	y_n'	$0.2y_n'$	$\frac{x_n + y_n}{2} + y_n' + 2$	$0.2y_n''$	$0.02y_n''$	$0.2y_n' + 0.02y_n''$
0	0	0	0	0	2.0000	0.0200	0.0200	0.0200
1	0.2	0.0200	0.2000	0.0400	2.4200	0.2420	0.0242	0.0642
2	0.4	0.0842	0.4420	0.0884	2.9262	0.2926	0.0293	0.1177
3	0.6	0.2019	0.7346	0.1469	3.5365	0.3537	0.0354	0.1823
4	0.8	0.3842	1.0883	0.2177	4.2725	0.4273	0.0427	0.2604
5	1.0	0.6446						

The student may show that the exact solution of the problem is

$$y = e^x - x - 1.$$

The errors of our approximate values are shown in Table 13 at the end of this section.

A much more accurate method is the **Runge-Kutta-Nyström method**¹⁴ which is a generalization of the Runge-Kutta method in Sec. 1.13. We mention without proof that this is a *fourth-order method*, which means that in the Taylor formulas for y and y' the first terms up to and including the term that contains h^4 are given exactly.

In the general step [the $(n + 1)$ th step] of the method, we first compute the auxiliary quantities

$$(6a) \quad \begin{aligned} A_n &= \frac{h}{2} f(x_n, y_n, y'_n) \\ B_n &= \frac{h}{2} f\left(x_n + \frac{h}{2}, y_n + \beta_n, y'_n + A_n\right) \quad \text{where } \beta_n = \frac{h}{2}\left(y'_n + \frac{A_n}{2}\right) \\ C_n &= \frac{h}{2} f\left(x_n + \frac{h}{2}, y_n + \beta_n, y'_n + B_n\right) \\ D_n &= \frac{h}{2} f(x_n + h, y_n + \delta_n, y'_n + 2C_n) \quad \text{where } \delta_n = h(y'_n + C_n) \end{aligned}$$

then the new value

$$(6b) \quad y_{n+1} = y_n + h(y'_n + K_n) \quad \text{where } K_n = \frac{1}{3}(A_n + B_n + C_n)$$

which is an approximation for $y(x_{n+1})$, and furthermore,

$$(6c) \quad y'_{n+1} = y'_n + K_n^* \quad \text{where } K_n^* = \frac{1}{3}(A_n + 2B_n + 2C_n + D_n)$$

which is an approximation for $y'(x_{n+1})$ needed in the next step.

Table 10 shows the arrangement of the computations of the first step.

Table 10. Runge-Kutta-Nyström Method, First Step.

n	x	y	y'	A_n	A_0
0	x_0	y_0	y'_0	A_0	A_0
	$x_0 + h/2$	$y_0 + \beta_0$	$y'_0 + A_0$	B_0	$2B_0$
	$x_0 + h/2$	$y_0 + \beta_0$	$y'_0 + B_0$	C_0	$2C_0$
	$x_0 + h$	$y_0 + \delta_0$	$y'_0 + 2C_0$		D_0
\dots				$3K_0$	$3K_0^*$
\dots				K_0	
				$h(y'_0 + K_0)$	K_0^*
		x_1	y_1		
			y'_1		

Example 2. Apply the Runge-Kutta-Nyström method to the initial value problem (5), choosing $h = 0.2$. Here $f = (x + y + y' + 2)/2$, and (6a) becomes

$$A_n = 0.05(x_n + y_n + y'_n + 2),$$

$$B_n = 0.05(x_n + 0.1 + y_n + \beta_n + y'_n + A_n + 2), \quad \beta_n = 0.1(y'_n + \frac{1}{2}A_n),$$

$$C_n = 0.05(x_n + 0.1 + y_n + \beta_n + y'_n + B_n + 2),$$

$$D_n = 0.05(x_n + 0.2 + y_n + \delta_n + y'_n + 2C_n + 2), \quad \delta_n = 0.2(y'_n + C_n).$$

¹⁴ E. J. NYSTRÖM, Finnish mathematician. His paper on the method appeared in *Acta Soc. Sci. Fennicae*, vol. 50 (1925).

Table 11 shows the computations of the first step, arranged as in Table 10.

Table 11. First Step of the Runge-Kutta-Nyström Method (with $h = 0.2$) Applied to the Initial Value Problem (5).

n	x	y	y'		
0	0	0	0	0.100 0000	0.100 0000
	0.1	0.005 0000	0.100 0000	0.110 2500	0.220 5000
	0.1	0.005 0000	0.110 2500	0.110 7625	0.221 5250
	0.2	0.022 1525	0.221 5250		0.122 1839
				0.321 0125	0.664 2089
				0.107 0042	
	0.2	0.021 4008	0.221 4030	0.021 4008	0.221 4030

In the present case the differential equation is simple, and so are the expressions for A_n , B_n , C_n , and D_n . Hence we may insert A_n into B_n , then B_n into C_n , and, finally, C_n into D_n . The result of this elementary calculation is

$$B_n = 0.05[1.0525(x_n + y_n) + 1.152 5y_n' + 2.205]$$

$$C_n = 0.05[1.055 125(x_n + y_n) + 1.160 125y_n' + 2.215 25]$$

$$D_n = 0.05[1.116 063 75(x_n + y_n) + 1.327 613 75y_n' + 2.443 677 5]$$

From this we may now determine K_n and K_n^* and insert the resulting expressions into (6b) and (6c), finding

$$(7) \quad \begin{aligned} y_{n+1} &= y_n + a(x_n + y_n) + by_n' + c \\ y_{n+1}' &= y_n' + a^*(x_n + y_n) + b^*y_n' + c^* \end{aligned}$$

where

$$\begin{aligned} a &= 0.010 3588 & b &= 0.211 0421 & c &= 0.021 4008 \\ : 0.105 5219 & : 0.115 8811 & : 0.221 4030 \end{aligned}$$

Table 12 shows the corresponding computations. The errors of the approximate values for $y(x)$ are much smaller than those in Ex. 1 (cf. Table 13).

Table 12. Runge-Kutta-Nyström Method (with $h = 0.2$) Applied to the Initial Value Problem (5); Five Steps Computed by the Use of (7).

n	x_n	y_n	y_n'	$a(x_n + y_n) + by_n' + c$	$a^*(x_n + y_n) + b^*y_n' + c^*$
0	0	0	0	0.021 4008	0.221 4030
1	0.2	0.021 4008	0.221 4030	0.070 4196	0.270 4220
2	0.4	0.091 8204	0.491 8250	0.130 2913	0.330 2940
3	0.6	0.222 1117	0.822 1190	0.203 4186	0.403 4219
4	0.8	0.425 5303	1.225 5409	0.292 7365	0.492 7403
5	1.0	0.718 2668	1.718 2812		

Table 13. Comparison of Accuracy of the Two Methods Under Consideration in the Case of the Initial Value Problem (5), with $h = 0.2$.

x	$y = e^x - x - 1$	Absolute value of error	
		Ex. 1	Table 12
0.2	0.021 4028	0.0014	0.000 0020
0.4	0.091 8247	0.0076	0.000 0043
0.6	0.222 1188	0.0202	0.000 0071
0.8	0.425 5409	0.0413	0.000 0106
1.0	0.718 2818	0.0737	0.000 0150

Problems

1. Repeat the computation in Ex. 1, choosing $h = 0.1$, and compare the errors of the values thus obtained with those in Ex. 1 (listed in Table 13).

Apply (4) to the following initial value problems. (Carry out five steps.)

2. $y'' = -y$, $y(0) = 0$, $y'(0) = 1$, $h = 0.1$
3. $y'' = -y$, $y(0) = 1$, $y'(0) = 0$, $h = 0.1$
4. $y'' = -y$, $y(0) = 1$, $y'(0) = 0$, $h = 0.05$
5. $y'' = y$, $y(0) = 1$, $y'(0) = 1$, $h = 0.1$
6. $y'' = y$, $y(0) = 1$, $y'(0) = -1$, $h = 0.1$

7. Show that in the case of the initial value problem

$$(1 - x^2)y'' - 2xy' + 2y = 0, \quad y(0) = 0, \quad y'(0) = 1$$

the formulas (4) with $h = 0.1$ take the form

$$y_{n+1} = y_n + 0.1y'_n + 0.01 \frac{x_n y_n' - y_n}{1 - x_n^2}$$

$$y'_{n+1} = y_n' + 0.2 \frac{x_n y_n' - y_n}{1 - x_n^2}$$

Carry out five steps. Verify by substitution that the exact solution of the problem is $y = x$.

8. Apply (4) with $h = 0.1$ to the initial value problem

$$(1 - x^2)y'' - 2xy' + 6y = 0, \quad y(0) = -\frac{1}{2}, \quad y'(0) = 0.$$

Carry out five steps. Verify by substitution that the exact solution is

$$y = \frac{1}{2}(3x^2 - 1).$$

9. Apply (4) with $h = 0.1$ to the initial value problem

$$y'' = xy' - 3y, \quad y(0) = 0, \quad y'(0) = -3.$$

Carry out five steps. Verify by substitution that the exact solution is $y = x^3 - 3x$.

10. Apply (4) with $h = 0.1$ to the initial value problem

$$y'' = xy' - 4y, \quad y(0) = 3, \quad y'(0) = 0.$$

Carry out five steps. Verify by substitution that the exact solution is

$$y = x^4 - 6x^2 + 3.$$

11. Apply the Runge-Kutta-Nyström method to the initial value problem in Prob. 2. Choose $h = 0.2$, carry out five steps, and compare the results with the exact values (cf. Table 1 in Sec. 0.1).

12. In Prob. 11, replace $h = 0.2$ by $h = 0.1$, carry out four steps, and compare the results with the corresponding values in Prob. 11 and with the exact values

$$0.099\,833\,417, \quad 0.198\,669\,331, \quad 0.295\,520\,207, \quad 0.389\,418\,342.$$

Chapter 3

Power Series Solutions of Differential Equations

In the preceding chapter we have seen that linear homogeneous differential equations with constant coefficients can be solved by algebraic methods, and the solutions are elementary functions known from calculus. In the case of equations with variable coefficients, the situation is more complicated and the solutions may be nonelementary functions. Bessel's equation, Legendre's equation, and the hypergeometric equation are of this type. Since these and other equations and their solutions play an important role in engineering mathematics, we shall now consider a method for solving such equations. The solutions will appear in the form of power series, and the method is, therefore, known as the power series method. We shall also consider some basic properties of the solutions, so that the student may get acquainted with these "higher transcendental functions" and with some standard procedures used in connection with special functions.

Prerequisite for this chapter: Chap. 2.

Sections which may be omitted in a shorter course: 3.2, 3.6, 3.7.

References: Appendix 1, Part B.

Answers to problems: Appendix 2.

✓ 3.1 *The Power Series Method*

We shall now consider solving differential equations by the so-called *power series method* which yields solutions in the form of power series. It is a very effective standard procedure in connection with linear differential equations whose coefficients are variable.

We first remember that a **power series**¹ (in powers of $x - a$) is an infinite series of the form

$$(1) \quad \sum_{m=0}^{\infty} c_m(x - a)^m = c_0 + c_1(x - a) + c_2(x - a)^2 + \cdots,$$

¹ The term "power series" alone usually refers to a series of the form (1), but does not include series of negative powers of x such as $c_0 + c_1x^{-1} + c_2x^{-2} + \cdots$ or series involving fractional powers of x . Note that in (1) we write, for convenience, $(x - a)^0 = 1$, even when $x = a$.

where c_0, c_1, \dots are constants, called the **coefficients** of the series, a is a constant, called the **center**, and x is a variable.

If in particular $a = 0$, we obtain a *power series in powers of x*

$$(2) \quad \sum_{m=0}^{\infty} c_m x^m = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \dots$$

We shall assume in this section that all variables and constants are real.

Familiar examples of power series are the Maclaurin series

$$\frac{1}{1-x} = \sum_{m=0}^{\infty} x^m = 1 + x + x^2 + \dots \quad (|x| < 1, \text{geometric series}),$$

$$e^x = \sum_{m=0}^{\infty} \frac{x^m}{m!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots,$$

$$\cos x = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{(2m)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots,$$

$$\sin x = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m+1}}{(2m+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots.$$

The basic idea of the power series method for solving differential equations is very simple and natural. We shall describe the practical procedure and illustrate it by simple examples, postponing the mathematical justification of the method to the next section.

A differential equation being given, we first represent all given functions in the equation by power series in powers of x (or in powers of $x - a$, if solutions in the form of power series in powers of $x - a$ are wanted). Then we assume a solution in the form of a power series, say,

$$(3) \quad y = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \dots = \sum_{m=0}^{\infty} c_m x^m$$

and insert this series and the series obtained by termwise differentiation,

$$(a) \quad y' = c_1 + 2c_2 x + 3c_3 x^2 + \dots = \sum_{m=1}^{\infty} m c_m x^{m-1}$$

$$(4) \quad (b) \quad y'' = 2c_2 + 3 \cdot 2c_3 x + 4 \cdot 3c_4 x^2 + \dots = \sum_{m=2}^{\infty} m(m-1)c_m x^{m-2}$$

etc., into the equation. Adding all the terms containing the same power of x , we may write the resulting equation in the form

$$(5) \quad k_0 + k_1 x + k_2 x^2 + \dots = 0$$

where the constants k_0, k_1, \dots are expressions containing the unknown coefficients c_0, c_1, \dots in (3). In order that (5) holds for all x in some interval, we must have

$$k_0 = 0, \quad k_1 = 0, \quad k_2 = 0, \dots$$

From these equations we may then determine the coefficients c_0, c_1, \dots successively.

Let us now illustrate the practical procedure for some simple equations which can also be solved by elementary methods.

Example 1. Solve

$$y' - y = 0.$$

In the first step, we insert (3) and (4a) into the equation:

$$(c_1 + 2c_2x + 3c_3x^2 + \dots) - (c_0 + c_1x + c_2x^2 + \dots) = 0.$$

Then we collect like powers of x , finding

$$(c_1 - c_0) + (2c_2 - c_1)x + (3c_3 - c_2)x^2 + \dots = 0.$$

Equating the coefficient of each power of x to zero, we have

$$c_1 - c_0 = 0, \quad 2c_2 - c_1 = 0, \quad 3c_3 - c_2 = 0, \dots.$$

Solving these equations, we may express c_1, c_3, \dots in terms of c_0 which remains arbitrary:

$$c_1 = c_0, \quad c_2 = \frac{c_1}{2} = \frac{c_0}{2!}, \quad c_3 = \frac{c_2}{3} = \frac{c_0}{3!},$$

With these values (3) becomes

$$y = c_0 + c_0x + \frac{c_0}{2!}x^2 + \frac{c_0}{3!}x^3 + \dots,$$

and we see that we have obtained the familiar general solution

$$y = c_0 \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \right) = c_0 e^x.$$

Example 2. Solve

$$y'' + y = 0.$$

By inserting (3) and (4b) into the equation we obtain

$$(2c_2 + 3 \cdot 2c_3x + 4 \cdot 3c_4x^2 + \dots) + (c_0 + c_1x + c_2x^2 + \dots) = 0.$$

Collecting like powers of x , we find

$$(2c_2 + c_0) + (3 \cdot 2c_3 + c_1)x + (4 \cdot 3c_4 + c_2)x^2 + \dots = 0.$$

Equating the coefficient of each power of x to zero, we have

$$2c_2 + c_0 = 0, \quad 3 \cdot 2c_3 + c_1 = 0, \quad 4 \cdot 3c_4 + c_2 = 0, \dots.$$

Solving these equations, we see that c_2, c_4, \dots may be expressed in terms of c_0 , and c_3, c_5, \dots may be expressed in terms of c_1 :

$$c_2 = -\frac{c_0}{2!}, \quad c_3 = -\frac{c_1}{3!}, \quad c_4 = -\frac{c_2}{4 \cdot 3} = \frac{c_0}{4!}, \dots;$$

c_0 and c_1 are arbitrary. With these values (3) becomes

$$y = c_0 + c_1x - \frac{c_0}{2!}x^2 - \frac{c_1}{3!}x^3 + \frac{c_0}{4!}x^4 + \frac{c_1}{5!}x^5 + \dots.$$

This may be written

$$y = c_0 \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots \right) + c_1 \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \right)$$

and we recognize the familiar general solution

$$y = c_0 \cos x + c_1 \sin x.$$

Example 3. Solve

$$(x+1)y' - (x+2)y = 0.$$

By inserting (3) and (4a) into the equation we obtain

$$(x+1)(c_1 + 2c_2x + 3c_3x^2 + \dots) - (x+2)(c_0 + c_1x + c_2x^2 + \dots) = 0.$$

By performing the indicated multiplications we get

$$\begin{aligned} & c_1x + 2c_2x^2 + 3c_3x^3 + 4c_4x^4 + 5c_5x^5 + \dots + sc_sx^s + \dots \\ & + c_1 + 2c_2x + 3c_3x^2 + 4c_4x^3 + 5c_5x^4 + 6c_6x^5 + \dots + (s+1)c_{s+1}x^s + \dots \\ & - c_0x - c_1x^2 - c_2x^3 - c_3x^4 - c_4x^5 - \dots - c_{s-1}x^s - \dots \\ & - 2c_0 - 2c_1x - 2c_2x^2 - 2c_3x^3 - 2c_4x^4 - 2c_5x^5 - \dots - 2c_sx^s - \dots = 0. \end{aligned}$$

Since this must be an identity in x , the sum of the coefficients of each power of x must be zero; we thus obtain

$$(6) \quad (a) c_1 - 2c_0 = 0 \quad (b) 2c_2 - c_1 - c_0 = 0, \quad \text{etc.,}$$

and in general

$$(7) \quad sc_s + (s+1)c_{s+1} - c_{s-1} - 2c_s = 0.$$

By solving (6a) for c_1 and (7) for c_{s+1} we have

$$(8) \quad (a) c_1 = 2c_0, \quad (b) c_{s+1} = \frac{1}{s+1} [c_{s-1} + (2-s)c_s], \quad s = 1, 2, \dots$$

Formula (8b) is a recursion formula, from which we may now determine c_2, c_3, \dots successively; if we wish, we may arrange the calculation in tabular form:

s	c_{s-1}	$(2-s)c_s$	Sum	$s+1$	$c_{s+1} = \frac{\text{Sum}}{s+1}$	c_{s+1} in terms of c_0
						$c_1 = 2c_0$
1	c_0	c_1	$c_0 + c_1$	2	$\frac{c_0}{2} + \frac{c_1}{2}$	$c_2 = \frac{3}{2}c_0$
2	c_1	0	c_1	3	$\frac{c_1}{3}$	$c_3 = \frac{2}{3}c_0$
3	c_2	$-c_3$	$c_2 - c_3$	4	$\frac{c_2}{4} - \frac{c_3}{4}$	$c_4 = \frac{5}{24}c_0$
...

With these values (3) becomes

$$y = c_0(1 + 2x + \frac{3}{2}x^2 + \frac{5}{8}x^3 + \frac{5}{24}x^4 + \dots),$$

and the reader may verify that the explicitly written terms are the first few terms of the Maclaurin series of

$$y = c_0(1+x)e^x,$$

the general solution of the equation obtained by separating variables.

Problems

Apply the power series method to the following differential equations.

1. $y' = 3y$

2. $y' + y = 0$

3. $y' = ky$

4. $(1 - x)y' = y$

5. $(1 - x^2)y' = y$

6. $y' = 2xy$

7. $y'' = y$

8. $y'' + 4y = 0$

9. $y'' = 4y$

10. $y' - y = 2x - x^2$

(More problems of this type are included at the end of Sec. 3.2.)

3.2 Theoretical Basis of the Power Series Method

We have seen that the power series method yields solutions of differential equations in the form of power series. The solution y of a given equation is assumed in the form of a power series with undetermined coefficients, and the coefficients are determined successively by inserting that series and the series for the derivatives of y into the given equation.

The practical usefulness of the method becomes obvious if we remember that power series may be used for computing values of the solutions. Furthermore, many general properties of the solutions can be derived from their power series; this will be seen in our further consideration.

The power series method involves various operations on power series, for example, differentiation, addition, and multiplication of power series. Therefore, to justify the method we have to consider the theoretical basis of these operations. This will be done in the present section. The discussion will involve some concepts and facts which may be already known to the student from elementary calculus, as well as other concepts which are not considered in elementary classes. [Corresponding proofs and further details (which will not be used in this chapter) can be found in Chap. 13.]

A power series is an infinite series of the form

$$(1) \quad \sum_{m=0}^{\infty} c_m(x - a)^m = c_0 + c_1(x - a) + c_2(x - a)^2 +$$

and we assume that the variable x , the center a , and the coefficients c_0, c_1, \dots are real, as in the previous section.

The expression

$$(2) \quad s_n(x) = c_0 + c_1(x - a) + \dots + c_n(x - a)^n$$

is called the *n*th partial sum of the series (1). Clearly if we omit the terms of s_n from (1), the remaining expression is

$$(3) \quad R_n(x) = c_{n+1}(x - a)^{n+1} + c_{n+2}(x - a)^{n+2} + \dots,$$

and this expression is called the remainder of (1) after the term $c_n(x - a)^n$.

For example, in the case of the geometric series

$$1 + x + x^2 + \cdots + x^n + \cdots$$

we have

$$\begin{aligned} s_1 &= 1 + x, & R_1 &= x^3 + x^5 + x^7 + \cdots, \\ s_2 &= 1 + x + x^2, & R_2 &= x^4 + x^6 + x^8 + \cdots, \text{ etc.} \end{aligned}$$

In this way we have now associated with (1) the sequence of the partial sums $s_1(x), s_2(x), \dots$. It may happen that for some $x = x_0$ this sequence converges, say,

$$\lim_{n \rightarrow \infty} s_n(x_0) = s(x_0).$$

Then we say that the series (1) converges, or is convergent, at $x = x_0$; the number $s(x_0)$ is called the value or sum of (1) at x_0 , and we write

$$s(x_0) = \sum_{m=0}^{\infty} c_m (x_0 - a)^m.$$

If that sequence is divergent at $x = x_0$, then the series (1) is said to diverge, or to be divergent, at $x = x_0$.

We remember that a sequence s_1, s_2, \dots is said to converge to a number s , or to be convergent with the limit s , if to each given positive number ϵ (no matter how small, but not zero) we can find a number N such that

$$(4) \quad |s_n - s| < \epsilon \quad \text{for each } n > N.$$

Geometrically speaking, (4) means that s_n with $n > N$ lies between $s - \epsilon$ and $s + \epsilon$ (Fig. 76). Of course, N will in general depend on the choice of ϵ .

Now in our case, $s = s_n + R_n$ or $R_n = s - s_n$. Hence

$$|s_n - s| = |R_n|$$

in (4), and convergence at $x = x_0$ means that we can make $|R_n(x_0)|$ as small as we please, by taking n large enough. In other words, in the case of convergence, $s_n(x_0)$ is an approximation of $s(x_0)$, and the error $|R_n(x_0)|$ of the approximation can be made smaller than any prescribed positive number ϵ by taking n sufficiently large.

If we choose $x = x_0 = a$ in (1), the series reduces to the single term c_0 because the other terms are zero. This shows that the series (1) converges at $x = a$. In some cases this may be the only value of x for which (1) converges. If there are other values of x for which the series converges, these values form an interval, the convergence interval, having the midpoint $x = a$.



Fig. 76. Convergence of a sequence.

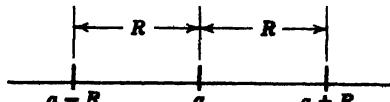


Fig. 77. Convergence interval (5) of a power series with center a.

This interval may be finite, as in Fig. 77. Then the series converges for all x in the interior of the interval, that is, for all x for which

$$(5) \quad |x - a| < R$$

and diverges when $|x - a| > R$. The interval may also be infinite, that is, the series may converge for all x .

The quantity R in Fig. 77 is called the **radius of convergence** of (1); it is the distance of each end point of the convergence interval from the center a . If the series converges for all x , then we set $R = \infty$ (and $1/R = 0$).

The radius of convergence can be determined from the coefficients of the series by means of each of the formulas

$$(6) \quad (a) \frac{1}{R} = \lim_{m \rightarrow \infty} \sqrt[m]{|c_m|} \quad (b) \frac{1}{R} = \lim_{m \rightarrow \infty} \left| \frac{c_{m+1}}{c_m} \right|$$

provided the limits in (6) exist. (Proofs of these facts can be found in Sec. 13.5.)

For each x for which (1) converges, it has a certain value $s(x)$ depending on x ; if the radius of convergence R of (1) is not zero, we write

$$s(x) = \sum_{m=0}^{\infty} c_m (x - a)^m \quad (|x - a| < R)$$

and we say that the series (1) *represents* the function $s(x)$ in the interval of convergence.

Example 1. In the case of the geometric series,

$$\frac{1}{1-x} = \sum_{m=0}^{\infty} x^m = 1 + x + x^2 + \dots \quad (|x| < 1).$$

In fact, $c_m = 1$ for all m , and from (6) we obtain $R = 1$, that is, the geometric series converges and represents $1/(1-x)$ when $|x| < 1$.

Example 2. In the case of the series

$$e^x = \sum_{m=0}^{\infty} \frac{x^m}{m!} = 1 + x + \frac{x^2}{2!} + \dots$$

we have $c_m' = 1/m!$. Hence in (6b),

$$\frac{c_{m+1}}{c_m} = \frac{1/(m+1)!}{1/m!} = \frac{1}{m+1} \rightarrow 0 \quad \text{as } m \rightarrow \infty,$$

which means $R = \infty$; the series converges for all x .

Example 3. In the case of the series

$$\sum_{m=0}^{\infty} m! x^m = 1 + x + 2x^2 + 6x^3 + \dots$$

we have $c_m = m!$, and in (6b),

$$\frac{c_{m+1}}{c_m} = \frac{(m+1)!}{m!} = m+1 \rightarrow \infty \quad \text{as } m \rightarrow \infty.$$

Thus $R = 0$, and the series converges only at the center $x = 0$.

We shall now consider the operations on power series which are used in connection with the power series method.

A power series may be differentiated term by term. More precisely: if

$$y(x) = \sum_{m=0}^{\infty} c_m (x - a)^m$$

converges for $|x - a| < R$, where $R > 0$, then the series obtained by differentiating term by term also converges for those x and represents the derivative y' of y , that is

$$y'(x) = \sum_{m=1}^{\infty} m c_m (x - a)^{m-1}.$$

(Proof in Sec. 13.6, Theorems 3 and 5.)

Two power series may be added term by term. More precisely: if the series

$$(7) \quad \sum_{m=0}^{\infty} b_m (x - a)^m \quad \text{and} \quad \sum_{m=0}^{\infty} c_m (x - a)^m$$

have positive radii of convergence and their sums are $f(x)$ and $g(x)$, then the series

$$\sum_{m=0}^{\infty} (b_m + c_m) (x - a)^m$$

converges and represents $f(x) + g(x)$ for each x which lies in the interior of the convergence interval of each of the given series (Proof in Sec. 13.5).

Two power series may be multiplied term by term. More precisely: Suppose that the series (7) have positive radii of convergence and let $f(x)$ and $g(x)$ be their sums. Then the series obtained by multiplying each term of the first series by each term of the second series and collecting like powers of $x - a$, that is,

$$b_0 c_0 + (b_0 c_1 + b_1 c_0)(x - a) + \cdots = \sum_{n=0}^{\infty} (b_0 c_n + b_1 c_{n-1} + \cdots + b_n c_0) (x - a)^n$$

converges and represents $f(x)g(x)$ for each x in the interior of the convergence interval of each of the given series. (Proof in Sec. 13.5, Theorem 3.)

If a power series has a positive radius of convergence and a sum which is identically zero throughout its interval of convergence, then each coefficient of the series is zero. (Proof in Sec. 13.6, Theorem 2.)

These properties of power series form the theoretical basis of the power series method. The remaining question is whether a given differential equation has solutions representable by power series at all. We may answer this question by using the following concept.

A function $f(x)$ is said to be *analytic at a point $x = a$* if it can be represented by a power series in powers of $x - a$ with radius of convergence $R > 0$.

Using this notion, a basic criterion of the desired type may be formulated as follows.

Theorem 1. *If the functions f , g , and r in the differential equation*

$$(8) \quad y'' + f(x)y' + g(x)y = r(x)$$

are analytic at $x = a$, then every solution $y(x)$ of (8) is analytic at $x = a$ and can thus be represented by a power series in powers of $x - a$ with radius of convergence² $R > 0$.

The proof requires advanced methods of complex analysis and can be found in Ref. [B5] in Appendix 1. In applying this theorem it is important to write the linear equation in the form (8), with 1 as the coefficient of y'' .

Problems

Determine the radius of convergence of the following series.

$$1. \sum_{m=0}^{\infty} \frac{x^m}{3^m}$$

$$2. \sum_{m=0}^{\infty} \frac{(-1)^m x^m}{m!}$$

$$3. \sum_{m=0}^{\infty} x^{2m}$$

$$4. \sum_{m=0}^{\infty} (-1)^m x^{3m}$$

$$5. \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{(2m)!}$$

$$6. \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m+1}}{(2m+1)!}$$

$$7. \sum_{m=1}^{\infty} \frac{x^m}{m}$$

$$8. \sum_{m=0}^{\infty} \frac{x^{2m+1}}{2m+1}$$

$$9. \sum_{m=0}^{\infty} \frac{x^{2m+1}}{(2m+1)m!}$$

$$10. \sum_{m=1}^{\infty} \frac{(4m)!}{(m!)^4} x^m$$

$$11. \sum_{m=1}^{\infty} \frac{(-1)^m (2m)!}{(m!)^2} x^m$$

$$12. \sum_{m=1}^{\infty} \frac{x^m}{2^m m^2}$$

$$13. \sum_{m=0}^{\infty} (-1)^m (x-1)^m$$

$$14. \sum_{m=0}^{\infty} \frac{1}{2^m} (x-3)^{2m}$$

$$15. \sum_{m=0}^{\infty} \frac{m}{3^m} (x-2)^m$$

Apply the power series method to the following differential equations.

$$16. xy' = (x+1)y$$

$$17. xy' - (x+2)y = -2x^2 - 2x$$

$$18. xy'' - 3y = 3$$

$$19. x(x+1)y' - (2x+1)y = 0$$

$$20. y'' - y = x$$

$$21. y'' - 3y' + 2y = 0$$

$$22. y'' - 4xy' + (4x^2 - 2)y = 0$$

$$23. (1-x^2)y'' - 2xy' + 2y = 0$$

$$24. (1-x^2)y'' - 2xy' + 6y = 0$$

$$25. y'' - xy' + y = 0$$

26. Show that $xy'' + y' + xy = 0$ does not satisfy the assumptions in Theorem 1, but has a power series in powers of x as a solution. Does this contradict Theorem 1?

27. Show that $y' = (y/x) + 1$ cannot be solved for y as a power series in x . Solve this equation for y as a power series in powers of $x - 1$. (Hint: introduce $t = x - 1$ as a new independent variable and solve the resulting equation for y as a power

² R is at least equal to the distance between the point $x = a$ and that point (or those points) closest to $x = a$ at which one of the functions f , g , r , as functions of a complex variable, is not analytic. (Note that that point may not lie on the x -axis, but somewhere in the complex plane.)

series in 1.) Compare the result with that obtained by the appropriate elementary method.

Solve for y as a power series in powers of $x - 1$:

$$28. \quad y' = 2y$$

$$29. \quad y'' - y = 0$$

$$30. \quad xy' + y = 0$$

3.3 Legendre's Equation. Legendre Polynomials

Legendre's differential equation³

$$(1) \quad (1 - x^2)y'' - 2xy' + n(n + 1)y = 0$$

arises in numerous physical problems, particularly in boundary value problems for spheres. The parameter n in (1) is a given real number. Any solution of (1) is called a **Legendre function**. We note that (1) may be written

$$(1') \quad [(1 - x^2)y']' + n(n + 1)y = 0.$$

Dividing (1) by $1 - x^2$, we obtain the standard form (8), Sec. 3.2, and we see that the coefficients of the resulting equation are analytic at $x = 0$, so that we may apply the power series method. Substituting

$$(2) \quad y = \sum_{m=0}^{\infty} c_m x^m$$

and its derivatives into (1) and denoting the constant $n(n + 1)$ by k we obtain

$$(1 - x^2) \sum_{m=2}^{\infty} m(m - 1)c_m x^{m-2} - 2x \sum_{m=1}^{\infty} mc_m x^{m-1} + k \sum_{m=0}^{\infty} c_m x^m = 0$$

or by writing the first expression as two separate series

$$\sum_{m=2}^{\infty} m(m - 1)c_m x^{m-2} - \sum_{m=2}^{\infty} m(m - 1)c_m x^m - 2 \sum_{m=1}^{\infty} mc_m x^m + k \sum_{m=0}^{\infty} c_m x^m = 0.$$

If we write each series at length, this becomes

$$\begin{aligned} 2 \cdot 1c_2 + 3 \cdot 2c_3x + 4 \cdot 3c_4x^2 + \cdots + (s+2)(s+1)c_{s+2}x^s &+ \cdots \\ - 2 \cdot 1c_2x^2 - \cdots &- s(s-1)c_sx^s - \cdots \\ - 2 \cdot 1c_1x - 2 \cdot 2c_2x^2 - \cdots &- 2sc_sx^s - \cdots \\ + kc_0 + kc_1x + kc_2x^2 + \cdots &+ kc_sx^s + \cdots = 0. \end{aligned}$$

³ADRIEN MARIE LEGENDRE (1752–1833), French mathematician, who made important contributions to the theory of numbers and elliptic functions.

Since this must be an identity in x if (2) is to be a solution of (1), the sum of the coefficients of each power of x must be zero; remembering that $k = n(n + 1)$ we thus have

$$(3a) \quad 2c_2 + n(n + 1)c_0 = 0, \quad 6c_3 + [-2 + n(n + 1)]c_1 = 0,$$

and in general, when $s = 2, 3, \dots$,

$$(3b) \quad (s + 2)(s + 1)c_{s+2} + [-s(s - 1) - 2s + n(n + 1)]c_s = 0.$$

The expression in brackets $[\dots]$ can be written

$$(n - s)(n + s + 1).$$

We thus obtain from (3)

$$(4) \quad c_{s+2} = -\frac{(n - s)(n + s + 1)}{(s + 2)(s + 1)} c_s \quad (s = 0, 1, \dots).$$

This is a recursion formula, giving each coefficient in terms of the second preceding it, except for c_0 and c_1 , which are left as arbitrary constants. find successively

$$\begin{array}{ll} c_2 = -\frac{n(n + 1)}{2!} c_0 & c_3 = -\frac{(n - 1)(n + 2)}{3!} c_1 \\ c_4 = -\frac{(n - 2)(n + 3)}{4 \cdot 3} c_2 & c_5 = -\frac{(n - 3)(n + 4)}{5 \cdot 4} c_3 \\ = \frac{(n - 2)n(n + 1)(n + 3)}{4!} c_0 & = \frac{(n - 3)(n - 1)(n + 2)(n + 4)}{5!} \end{array}$$

etc. By inserting these values for the coefficients into (2) we obtain

$$(5) \quad y(x) = c_0 y_1(x) + c_1 y_2(x)$$

where

$$(6) \quad y_1(x) = 1 - \frac{n(n + 1)}{2!} x^2 + \frac{(n - 2)n(n + 1)(n + 3)}{4!} x^4 - + \dots$$

and

$$(7) \quad y_2(x) = x - \frac{(n - 1)(n + 2)}{3!} x^3 + \frac{(n - 3)(n - 1)(n + 2)(n + 4)}{5!} x^5 - + \dots$$

These series converge for $|x| < 1$. Since (6) contains even powers of x only, while (7) contains odd powers of x only, the ratio y_1/y_2 is not a constant, and y_1 and y_2 are linearly independent solutions. Hence (5) is a general solution of (1) on the interval $-1 < x < 1$.

In many applications the parameter n in Legendre's equation will be a nonnegative integer. Then the right hand side of (4) is zero when $s = n$, and, therefore, $c_{n+2} = 0, c_{n+4} = 0, c_{n+6} = 0, \dots$. Hence, if n is even, $y_1(x)$ reduces to a polynomial of degree n . If n is odd, the same is true with respect to $y_2(x)$. These polynomials, multiplied by some constants, are called

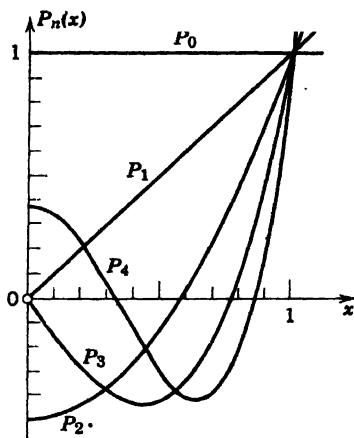


Fig. 78. Legendre polynomials.

Legendre polynomials. Since they are of great practical importance, let us consider them in more detail. For this purpose we write (4) in the form

$$(8) \quad c_s = -\frac{(s+2)(s+1)}{(n-s)(n+s+1)} c_{s+2} \quad (s \leq n-2)$$

and may then express all the nonvanishing coefficients in terms of the coefficient c_n of the highest power of x of the polynomial. The coefficient c_n is then arbitrary. It is customary to choose $c_n = 1$ when $n = 0$ and

$$(9) \quad c_n = \frac{(2n)!}{2^n(n!)^2} = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{n!}, \quad n = 1, 2, \dots,$$

the reason being that for this choice of c_n all those polynomials will have the value 1 when $x = 1$; this follows from (17) in Prob. 18. We then obtain from (8) and (9)

$$(9^*) \quad \begin{aligned} c_{n-2} &= -\frac{n(n-1)}{2(2n-1)} c_n = -\frac{n(n-1)(2n)!}{2(2n-1)2^n(n!)^2} \\ &= -\frac{n(n-1)2n(2n-1)(2n-2)!}{2(2n-1)2^n n(n-1)! n(n-1)(n-2)!}, \end{aligned}$$

that is,

$$c_{n-2} = -\frac{(2n-2)!}{2^n(n-1)!(n-2)!}.$$

Similarly,

$$c_{n-4} = -\frac{(n-2)(n-3)}{4(2n-3)} c_{n-2} = \frac{(2n-4)!}{2^n 2! (n-2)! (n-4)!}$$

etc., and in general, when $n-2m \geq 0$,

$$(10) \quad c_{n-2m} = (-1)^m \frac{(2n-2m)!}{2^m m! (n-m)! (n-2m)!}.$$

The resulting solution of Legendre's equation is called the **Legendre polynomial of degree n** and is denoted by $P_n(x)$; from (10) we obtain

$$(11) \quad \begin{aligned} P_n(x) &= \sum_{m=0}^M (-1)^m \frac{(2n-2m)!}{2^m m! (n-m)! (n-2m)!} x^{n-2m} \\ &\quad \frac{(2n)!}{2^n (n!)^2} x^n - \frac{(2n-2)!}{2^n 1! (n-1)! (n-2)!} x^{n-2} + \end{aligned}$$

where $M = n/2$ or $(n - 1)/2$, whichever is an integer. In particular (Fig. 78),

$$(11') \quad \begin{aligned} P_0(x) &= 1, & P_1(x) &= x, \\ P_2(x) &= \frac{1}{2}(3x^2 - 1), & P_3(x) &= \frac{1}{2}(5x^3 - 3x), \\ P_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3), & P_5(x) &= \frac{1}{8}(63x^5 - 70x^3 + 15x), \\ \text{etc.} & & & \end{aligned}$$

Problems

1. Derive (11') from (11). Find and graph $P_6(x)$.
2. Verify by inserting that $P_n(x)$, $n = 0, 1, \dots, 5$ [cf. (11')] is a solution of (1).
3. Show that the equation in Prob. 23, Sec. 3.2, is a special Legendre equation and obtain a general solution from (5), this section.
4. Find the radius of convergence of the series (6) and (7).
5. Using (5)–(7), find a general solution of the Legendre equation in Prob. 24, Sec. 3.2.
6. For what initial values does the solution in Prob. 5 yield $P_2(x)$?
7. Show that $P_n(-x) = (-1)^n P_n(x)$ and $P_n'(-x) = (-1)^{n+1} P_n'(x)$.
8. (Orthogonality of the Legendre polynomials). The important property expressed by the formula

$$(12) \quad \int_{-1}^1 P_m(x) P_n(x) dx = 0 \quad \text{when } m \neq n$$

is called the *orthogonality of the Legendre polynomials on the interval $-1 \leq x \leq 1$* . (This property will be considered from a more general point of view in Sec. 8.12.) Verify (12) for $m, n = 0, 1, 2$ by direct calculation.

9. Prove (12) as follows. Start from the Legendre equations

$$[(1 - x^2)P_m]' = -m(m + 1)P_m, \quad [(1 - x^2)P_n]' = -n(n + 1)P_n.$$

Multiply the first by P_n , the last by $-P_m$, and add. Integrate the resulting equation over x from -1 to 1 . Integrating the integrals on the left by parts, show that they are zero and this implies (12).

10. Prove

$$(13) \quad P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n] \quad (\text{Rodrigues' formula})$$

by applying the binomial theorem to $(x^2 - 1)^n$, differentiating the result n times term by term, and comparing with (11).

11. Using (13) and integrating n times by parts, show that

$$(14) \quad \int_{-1}^1 P_n^2(x) dx = \frac{2}{2n + 1} \quad (n = 0, 1, \dots).$$

12. Using induction, show that any polynomial $f(x)$ of degree n can be represented in the form

$$(15) \quad f(x) = c_0 P_0 + c_1 P_1(x) + \cdots + c_n P_n(x).$$

13. Show that in (15),

$$(16) \quad c_m = \frac{2m + 1}{2} \int_{-1}^1 f(x) P_m(x) dx \quad (m = 0, 1, \dots, n).$$

Represent the following polynomials in the form (15).

14. $1 - 3x + 3x^2$

15. $2x + 10x^3$

16. $3x - 5x^3 + 70x^4$

17. $1, x, x^2, x^3$

18. Show that

$$(17) \quad \frac{1}{\sqrt{1 - 2xu + u^2}} = \sum_{n=0}^{\infty} P_n(x)u^n.$$

Hint: start from the binomial expansion of $1/\sqrt{1 - v}$, set $v = 2xu - u^2$, multiply the powers of $2xu - u^2$ out, collect all the terms involving u^n , and verify that the sum of these terms is $P_n(x)u^n$.

19. Let A_1 and A_2 be two points in space (Fig. 79, $r_2 > 0$). Using (17), show that

$$\frac{1}{r} = \frac{1}{\sqrt{r_1^2 + r_2^2 - 2r_1r_2 \cos \theta}} = \frac{1}{r_2} \left[P_0 + P_1(\cos \theta) \frac{r_1}{r_2} + P_2(\cos \theta) \left(\frac{r_1}{r_2} \right)^2 + \dots \right].$$

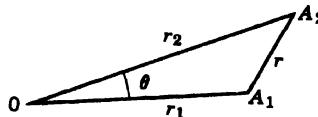


Fig. 79. Problem 19.

20. Using the idea in Prob. 9 and the fact that $y_n = \cos nx$ satisfies the differential equation $y'' + n^2y = 0$, prove the orthogonality relation

$$\int_0^{2\pi} \cos mx \cos nx dx = 0 \quad (m \text{ and } n \text{ integral, } m \neq n).$$

3.4 Extended Power Series Method. Indicial Equation

Several second-order differential equations, which are highly important in many applications, have coefficients which are not analytic at $x = 0$ (definition in Sec. 3.2), but are such that the following theorem can be proved.

Theorem 1. Any differential equation of the form

$$(1) \quad y'' + \frac{a(x)}{x} y' + \frac{b(x)}{x^2} y = 0,$$

where the functions $a(x)$ and $b(x)$ are analytic at $x = 0$, has at least one solution which can be represented in the form

$$(2) \quad y(x) = x^r \sum_{m=0}^{\infty} c_m x^m = x^r (c_0 + c_1 x + c_2 x^2 + \dots) \quad (c_0 \neq 0)$$

where the exponent r may be any (real or complex) number (and is chosen so that $c_0 \neq 0$).⁴

⁴ In this theorem, the variable x may be replaced by $x - a$ where a is any number.

The proof requires advanced methods of complex analysis and can be found in Ref. [B5], Appendix 1. The following method for solving (1) is called the **Frobenius' method**.⁵

To solve (1), we write it in the somewhat more convenient form

$$(1') \quad x^2y'' + xa(x)y' + b(x)y = 0.$$

We first expand $a(x)$ and $b(x)$ in power series,

$$a(x) = a_0 + a_1x + a_2x^2 + \cdots, \quad b(x) = b_0 + b_1x + b_2x^2 + \cdots.$$

Then we differentiate (2) term by term, finding

$$\begin{aligned} y'(x) &= \sum_{m=0}^{\infty} (m+r)c_m x^{m+r-1} = x^{r-1}[rc_0 + (r+1)c_1x + \cdots], \\ y''(x) &= \sum_{m=0}^{\infty} (m+r)(m+r-1)c_m x^{m+r-2} \\ &= x^{r-2}[r(r-1)c_0 + (r+1)rc_1x + \cdots]. \end{aligned}$$

By inserting all these series into (1') we readily obtain

$$\begin{aligned} x^r[r(r-1)c_0 + \cdots] + (a_0 + a_1x + \cdots)x^r(c_0 + \cdots) \\ + (b_0 + b_1x + \cdots)x^r(c_0 + c_1x + \cdots) = 0. \end{aligned}$$

We now equate the sum of the coefficients of each power of x to zero, as before. This yields a system of equations involving the unknown coefficients c_m . The smallest power is x^r , and the corresponding equation is

$$[r(r-1) + a_0r + b_0]c_0 = 0.$$

Since by assumption $c_0 \neq 0$, we obtain

$$(3) \quad r^2 + (a_0 - 1)r + b_0 = 0.$$

This important quadratic equation is called the **indicial equation** of the differential equation (1). We shall see that our method will yield a fundamental system of solutions; one of the solutions will always be of the form (2), but for the form of the other solution there will be three different possibilities corresponding to the following cases.

Case 1. *The roots of the indicial equation are distinct and do not differ by an integer ($1, 2, 3, \dots$)*.⁶

Case 2. *The indicial equation has a double root.*

Case 3. *The roots of the indicial equation differ by an integer.*

⁵ GEORG FROBENIUS (1849–1917), German mathematician, who also made important contributions to the theory of matrices and groups.

⁶ Note that this case includes complex conjugate roots r_1 and $r_2 = \bar{r}_1$, because $r_1 - r_2 = r_1 - \bar{r}_1 = 2i \operatorname{Im} r_1 \neq \pm 1, \pm 2, \dots$.

Let us discuss these cases separately.

Case 1. Distinct roots of the indicial equation not differing by an integer.

This is the simplest case. Let r_1 and r_2 be the roots of (3). If we insert $r = r_1$ into the aforementioned system of equations and determine the coefficients c_1, c_2, \dots successively, as before, then we obtain a solution

$$(4) \quad y_1(x) = x^{r_1}(c_0 + c_1x + c_2x^2 + \dots).$$

It can be shown (cf. Ref. [B5]) that in the present case the differential equation has another independent solution $y_2(x)$ of a similar form, namely

$$(5) \quad y_2(x) = x^{r_2}(c_0^* + c_1^*x + c_2^*x^2 + \dots),$$

which can be obtained by inserting $r = r_2$ into that system of equations and determining the coefficients c_0^*, c_1^*, \dots successively. Linear independence of y_1 and y_2 follows from the fact that y_1/y_2 is not constant because $r_1 - r_2$ is not an integer. Hence y_1 and y_2 form a fundamental system of (1) on the interval of convergence of both series.

Example 1 (Case 1). We consider the differential equation

$$x^2y'' + (x^2 + \frac{5}{36})y = 0.$$

By substituting (2) into this equation we obtain

$$(6) \quad \sum_{m=0}^{\infty} (m+r)(m+r-1)c_m x^{m+1} + \sum_{m=0}^{\infty} c_m x^{m+r+2} + \frac{5}{36} \sum_{m=0}^{\infty} c_m x^{m+r} = 0.$$

By equating the sum of the coefficients of x^r to zero we obtain the indicial equation

$$r(r-1) + \frac{5}{36} = 0.$$

The roots are $r_1 = 5/6$ and $r_2 = 1/6$. By equating the sum of the coefficients of x^{r+s} in (6) to zero, we find

$$(7) \quad \begin{aligned} (a) \quad & [(r+1)r + \frac{5}{36}]c_1 = 0 & (s = 1) \\ (b) \quad & (s+r)(s+r-1)c_s + c_{s-2} + \frac{5}{36}c_s = 0 & (s = 2, 3, \dots). \end{aligned}$$

Let us first determine a solution $y_1(x)$ corresponding to $r_1 = 5/6$. For this value of r , equation (7a) yields $c_1 = 0$. Equation (7b) may be written

$$(8) \quad s(s + \frac{2}{3})c_s + c_{s-2} = 0.$$

From $c_1 = 0$ it follows that $c_3 = 0, c_5 = 0, \dots$, successively. If we set $s = 2p$ in (8), we obtain

$$c_{2p} = -\frac{3}{4} \cdot \frac{c_{2p-2}}{p(3p+1)} \quad (p = 1, 2, \dots).$$

From this formula we see that the coefficients are

$$c_2 = -\frac{3}{4} \cdot \frac{c_0}{4}$$

$$c_4 = -\frac{3}{4} \cdot \frac{c_2}{2 \cdot 7} = \left(\frac{3}{4}\right)^2 \frac{c_0}{2! \cdot 4 \cdot 7}$$

$$c_6 = -\frac{3}{4} \cdot \frac{c_4}{3 \cdot 10} = -\left(\frac{3}{4}\right)^3 \frac{c_0}{3! \cdot 4 \cdot 7 \cdot 10}, \quad \text{etc.}$$

This yields the solution

$$(9) \quad y_1(x) = c_0 \sum_{p=0}^{\infty} (-1)^p \left(\frac{3}{4}\right)^p \frac{x^{2p+\frac{1}{4}}}{p! 1 \cdot 4 \cdots (3p+1)} = c_0 x^{\frac{1}{4}} \left(1 - \frac{3}{16}x^2 + \frac{9}{896}x^4 - + \cdots\right).$$

Another solution $y_2(x)$ corresponding to the root $r_2 = \frac{1}{3}$ may be obtained in a similar fashion. Denoting the coefficients of this solution by c_0^*, c_1^*, \dots , instead of (7) we now have

$$(10) \quad \begin{aligned} (a) \quad & [(r_2 + 1)r_2 + \frac{5}{3}]c_1^* = 0 \\ (b) \quad & (s + r_2)(s + r_2 - 1)c_s^* + c_{s-2}^* + \frac{5}{3}c_s^* = 0 \quad (s = 2, 3, \dots). \end{aligned}$$

From (10a) we see that $c_1^* = 0$, and (10b) may be written

$$(11) \quad s(s - \frac{2}{3})c_s^* + c_{s-2}^* = 0 \quad (s = 2, 3, \dots).$$

Since $c_1^* = 0$, it follows that $c_3^* = 0, c_5^* = 0, \dots$, successively. If we set $s = 2p$ in (11), we obtain

$$c_{2p}^* = -\frac{3}{4} \cdot \frac{c_{2p-2}^*}{p(3p-1)} \quad (p = 1, 2, \dots).$$

From this formula we find

$$\begin{aligned} c_2^* &= -\frac{3}{4} \cdot \frac{c_0^*}{2} \\ c_4^* &= -\frac{3}{4} \cdot \frac{c_2^*}{2 \cdot 5} = \left(\frac{3}{4}\right)^2 \frac{c_0^*}{2! 2 \cdot 5}, \quad \text{etc.,} \end{aligned}$$

and the corresponding solution is

$$(12) \quad y_2(x) = c_0^* + c_0^* \sum_{p=1}^{\infty} (-1)^p \left(\frac{3}{4}\right)^p \frac{x^{2p+\frac{1}{4}}}{p! 2 \cdot 5 \cdots (3p-1)} = c_0^* x^{\frac{1}{4}} \left(1 - \frac{3}{8}x^2 + - \cdots\right).$$

The series (9) and (12) converge for all x , and y_1 and y_2 are linearly independent. Hence these functions constitute a fundamental system for all values of x . (In the next section we shall see that our equation is closely related to Bessel's equation with the parameter $\nu = 1/3$.)

Case 2. Double root of the indicial equation. The indicial equation (3) has a double root r if, and only if, $(a_0 - 1)^2 - 4b = 0$, and then $r = (1 - a_0)/2$. We may determine a first solution

$$(13) \quad y_1(x) = x^r(c_0 + c_1x + c_2x^2 + \cdots) \quad \left(r = \frac{1 - a_0}{2}\right)$$

as before. To find another solution we may apply the method of variation of parameters, that is, we replace the constant c in the solution cy_1 by a function $u(x)$ to be determined such that

$$(14) \quad y_2(x) = u(x)y_1(x)$$

is a solution of (1). By inserting (14) and the derivatives

$$y_1' = u'y_1 + uy_1', \quad y_2'' = u''y_1 + 2u'y_1' + uy_1''$$

into the differential equation (1') we obtain

$$x^a(u''y_1 + 2u'y_1' + uy_1'') + xa(u'y_1 + uy_1') + buy_1 = 0.$$

Since y_1 is a solution of (1'), the sum of the terms involving u is zero, and the last equation reduces to

$$x^2y_1u'' + 2x^2y_1'u' + xay_1u' = 0.$$

By dividing by x^2y_1 and inserting the power series for a we obtain

$$u'' + \left(2\frac{y_1'}{y_1} + \frac{a_0}{x} + \dots\right)u' = 0.$$

Here and in the following the dots designate terms which are constant or involve positive powers of x . Now from (13) it follows that

$$\begin{aligned}\frac{y_1'}{y_1} &= \frac{x^{r-1}[rc_0 + (r+1)c_1x + \dots]}{x^r[c_0 + c_1x + \dots]} \\ &= \frac{1}{x} \left(\frac{rc_0 + (r+1)c_1x + \dots}{c_0 + c_1x + \dots} \right) = \frac{r}{x} + \dots.\end{aligned}$$

Hence the last equation can be written

$$(15) \quad u'' + \left(\frac{2r + a_0}{x} + \dots\right)u' = 0.$$

Since $r = (1 - a_0)/2$ the term $(2r + a_0)/x$ equals $1/x$, and by dividing by u' we thus have

$$\frac{u''}{u'} = -\frac{1}{x} + \dots.$$

By integration we obtain

$$\ln u' = -\ln x + \dots \quad \text{or} \quad u' = \frac{1}{x} e^{(\dots)}.$$

Expanding the exponential function in powers of x and integrating once more, we see that the expression for u will be of the form

$$u = \ln x + k_1x + k_2x^2 + \dots$$

By inserting this into (14) we have the following result.

Form of the solutions in the case of a double root of the indicial equation. If the indicial equation (3) has a double root r , then linearly independent solutions of (1) are of the form (13) and

$$(16) \quad y_2(x) = y_1(x) \ln x + x^r \sum_{m=1}^{\infty} A_m x^m.$$

The simplest illustration of this case is given by the Cauchy equation (Sec. 2.7) in the critical case (cf. also Ex. 5, below).

Example 2 (Case 2). Solve the differential equation

$$(17) \quad x(x-1)y'' + (3x-1)y' + y = 0.$$

Writing this equation in the standard form (1), we see that it satisfies the assumptions in Theorem 1. By inserting (2) and its derivatives into (17) we have

$$(18) \quad \sum_{m=0}^{\infty} (m+r)(m+r-1)c_m x^{m+r} - \sum_{m=0}^{\infty} (m+r)(m+r-1)c_m x^{m+r-1} + 3 \sum_{m=0}^{\infty} (m+r)c_m x^{m+r} - \sum_{m=0}^{\infty} (m+r)c_m x^{m+r-1} + \sum_{m=0}^{\infty} c_m x^{m+r} = 0.$$

The smallest power is x^{r-1} ; by equating the sum of its coefficients to zero we have

$$[-r(r-1) - r]c_0 = 0 \quad \text{or} \quad r^2 = 0.$$

Hence this indicial equation has the double root $r = 0$. We insert this value into (18) and equate the sum of the coefficients of the power x^s to zero, finding

$$s(s-1)c_s - (s+1)sc_{s+1} + 3sc_s - (s+1)c_{s+1} + c_s = 0$$

or $c_{s+1} = c_s$. Hence $c_0 = c_1 = c_2 = \dots$, and by choosing $c_0 = 1$ we obtain

$$y_1(x) = \sum_{m=0}^{\infty} x^m = \frac{1}{1-x}.$$

To obtain a second solution we substitute $y_2 = uy_1$ into (17), finding

$$x(x-1)(u''y_1 + 2u'y_1' + uy_1'') + (3x-1)(u'y_1 + uy_1') + uy_1 = 0.$$

Since y_1 is a solution of (17), this reduces to

$$x(x-1)(u''y_1 + 2u'y_1') + (3x-1)u'y_1 = 0.$$

By inserting the expressions for y_1 and y_1' and simplifying we obtain

$$xu'' + u' = 0 \quad \text{or} \quad \frac{u''}{u'} = -\frac{1}{x}.$$

By integrating twice we finally obtain $u = \ln x$, and a second solution is

$$y_2 = uy_1 = \frac{\ln x}{1-x}.$$

Case 3. Roots of the indicial equation differing by an integer. If the roots r_1 and r_2 of the indicial equation (3) differ by an integer, say, $r_1 = r$ and $r_2 = r - p$, where p is a positive integer, then we may always determine one solution as before, namely, the solution

$$y_1(x) = x^{r_1}(c_0 + c_1x + c_2x^2 + \dots)$$

corresponding to the root r_1 . However, for the root r_2 there may arise difficulties, and it may not be possible to determine a solution y_2 by the method used in Case 1 (cf. Ex. 3, below). To determine y_2 , we may proceed as in Case 2. The first steps are literally the same and yield equation (15). We determine $2r + a_0$ in (15). From elementary algebra we know that the coefficient $a_0 - 1$ in (3) equals $-(r_1 + r_2)$. In our case, $r_1 = r$, $r_2 = r - p$, and, therefore, $a_0 - 1 = p - 2r$. Hence in (15), $2r + a_0 = p + 1$, and we obtain

$$\frac{u''}{u'} = -\left(\frac{p+1}{x} + \dots\right).$$

The further steps are as in Case 2. Integrating, we find

$$\ln u' = -(p+1) \ln x + \dots \quad \text{or} \quad u' = x^{-(p+1)} e^{(\dots)},$$

where the dots stand for some series of non-negative powers of x . By expanding the exponential function as before we obtain a series of the form

$$u' = \frac{1}{x^{p+1}} + \frac{k_1}{x^p} + \dots + \frac{k_p}{x} + k_{p+1} + k_{p+2}x +$$

By integrating once more we have

$$u = -\frac{1}{px^p} - \dots + k_p \ln x + k_{p+1}x + \dots.$$

Multiplying this expression by the series

$$y_1 = x^{r_1}(c_0 + c_1x + \dots)$$

and remembering that $r_1 - p = r_2$ we see that $y_2 = uy_1$ is of the form

$$(19) \quad y_2(x) = k_p y_1(x) \ln x + x^{r_2} \sum_{m=0}^{\infty} C_m x^m.$$

We may sum up our result as follows.

Form of the solutions when the roots of the indicial equation differ by an integer. If the roots of the indicial equation (3) differ by an integer, then to the root with the larger real part (say, r_1) there corresponds a solution of (1) of the form (4), and to the other root (r_2) there corresponds a solution of the form (19). The solutions form a fundamental system of (1).

While for a double root of (3) the second solution always contains a logarithmic term, the coefficient k_p in (19) may be zero and in the present case the logarithmic term may be missing (cf. Ex. 4).⁷

Example 3 (Case 3). We consider the differential equation

$$(x^2 - 1)x^2 y'' - (x^2 + 1)xy' + (x^2 + 1)y = 0.$$

By substituting (2) and its derivatives into the equation we first have

$$(x^2 - 1) \sum_{m=0}^{\infty} (m+r)(m+r-1)c_m x^{m+r} - (x^2 + 1) \sum_{m=0}^{\infty} (m+r)c_m x^{m+r} + (x^2 + 1) \sum_{m=0}^{\infty} c_m x^{m+r} = 0.$$

Performing the indicated multiplications and simplifying, we obtain

$$(20) \quad \sum_{m=0}^{\infty} (m+r-1)^2 c_m x^{m+r+2} - \sum_{m=0}^{\infty} (m+r+1)(m+r-1)c_m x^{m+r} = 0.$$

By equating the coefficient of x^r to zero we obtain the indicial equation

$$(r+1)(r-1) = 0.$$

⁷ A general consideration about the convergence of those series will not be presented here, but in each individual case convergence may be tested in the usual way.

The roots $r_1 = 1$ and $r_2 = -1$ differ by an integer. By equating the coefficient of x^{r+1} in (20) to zero we have

$$-(r + 2)c_1 = 0.$$

For $r = r_1$ as well as for $r = r_2$ this implies $c_1 = 0$. By equating the sum of the coefficients of x^{s+r+2} in (20) to zero we find

$$(21) \quad (s + r - 1)^2 c_s = (s + r + 3)(s + r + 1)c_{s+2} \quad (s = 0, 1, \dots).$$

Inserting $r = r_1 = 1$ and solving for c_{s+2} , we find

$$c_{s+2} = \frac{s^2}{(s + 4)(s + 2)} c_s \quad (s = 0, 1, \dots).$$

Since $c_1 = 0$, it follows that $c_3 = 0$, $c_5 = 0$, etc. For $s = 0$ we obtain $c_2 = 0$, and from this by taking $s = 2, 4, \dots$, we have $c_4 = 0$, $c_6 = 0$, etc. Hence to the larger root $r_1 = 1$ there corresponds the solution

$$y_1 = c_0 x.$$

If we insert the smaller root $r = r_2 = -1$ into (21), we get

$$(s - 2)^2 c_s = (s + 2) s c_{s+2} \quad (s = 0, 1, \dots).$$

With $s = 0$ this becomes $4c_0 = 0$, which requires $c_0 = 0$. This contradicts our initial assumption $c_0 \neq 0$. In fact, we shall now see that our equation does not have a second independent solution $y_2(x)$ of the form (2) corresponding to the smaller root r_2 . From (19) we see that $y_2(x)$ must be of the form

$$y_2(x) = kx \ln x + \frac{1}{x} \sum_{m=0}^{\infty} C_m x^m.$$

Substituting this and the derivatives into the differential equation, we obtain

$$(x^2 - 1)x^2 \left(\frac{k}{x} + \sum_{m=0}^{\infty} (m-1)(m-2)C_m x^{m-3} \right) - (x^2 + 1)x \left(k \ln x + k + \sum_{m=0}^{\infty} (m-1)C_m x^{m-2} \right) + (x^2 + 1) \left(kx \ln x + \sum_{m=0}^{\infty} C_m x^{m-1} \right) = 0.$$

The logarithmic terms drop out, and simplification yields

$$-2kx + \sum_{m=0}^{\infty} (m-2)^2 C_m x^{m+1} - \sum_{m=0}^{\infty} m(m-2)C_m x^{m-1} = 0.$$

By equating the sum of the coefficients of x^s to zero we obtain $C_1 = 0$ ($s = 0$), $-2k + 4C_0 = 0$ ($s = 1$), and

$$(s - 3)^2 C_{s-1} - (s^2 - 1)C_{s+1} = 0 \quad (s = 2, 3, \dots).$$

Solving for C_{s+1} , we have

$$C_{s+1} = \frac{(s - 3)^2}{s^2 - 1} C_{s-1} \quad (s = 2, 3, \dots).$$

Since $C_1 = 0$, we see that $C_3 = 0$, $C_5 = 0$, etc. $s = 3$ yields $C_4 = 0$; hence $C_6 = 0$, $C_8 = 0$, etc. $k = 2C_0$ and C_2 remain arbitrary. Consequently,

$$y_2(x) = 2C_0 x \ln x + \frac{1}{x} (C_0 + C_2 x^2).$$

Since the last term is $C_2 y_1/c_0$, we may take $C_2 = 0$. With $C_0 = 1/2$,

$$y_1(x) = x \ln x + \frac{1}{2x}.$$

In Cases 2 and 3 it may sometimes be simpler to obtain a second independent solution by the method of variation of parameters.

Let us illustrate this for the equation under consideration. Since x is a solution, we may set

$$y_2(x) = xu(x).$$

Substituting this into the equation and simplifying, we find

$$(x^3 - x)u'' + (x^2 - 3)u' = 0.$$

Using partial fractions, we obtain

$$\frac{u''}{u'} = \frac{3 - x^2}{x^3 - x} = -\frac{3}{x} + \frac{1}{x+1} + \frac{1}{x-1}.$$

Integrating on both sides, we have

$$\ln u' = -3 \ln x + \ln(x+1) + \ln(x-1) = \ln \frac{x^2 - 1}{x^3}.$$

Taking exponentials and integrating again, we find

$$u = \ln x + 1/(2x^3).$$

This yields the same expression as before, namely,

$$y_2(x) = xu(x) = x \ln x + \frac{1}{2x}.$$

Example 4 (Case 3, second solution without logarithmic term). We consider the differential equation

$$x^2 y'' + xy' + (x^2 - \frac{1}{4})y = 0.$$

Substituting (2) and the derivatives into this equation, we obtain

$$\sum_{m=0}^{\infty} [(m+r)(m+r-1) + (m+r) - \frac{1}{4}]c_m x^{m+r} + \sum_{m=0}^{\infty} c_m x^{m+r+2} = 0.$$

By equating the coefficient of x^r to zero we get the indicial equation

$$r(r-1) + r - \frac{1}{4} = 0 \quad \text{or} \quad r^2 = \frac{1}{4}.$$

The roots $r_1 = 1/2$ and $r_2 = -1/2$ differ by an integer. By equating the sum of the coefficients of x^{s+r} to zero we find

$$(22) \quad \begin{aligned} (a) \quad & [(r+1)r + (r+1) - \frac{1}{4}]c_1 = 0 & (s=1) \\ (b) \quad & [(s+r)(s+r-1) + s+r - \frac{1}{4}]c_s + c_{s-2} = 0 & (s=2, 3, \dots). \end{aligned}$$

For $r = r_1 = 1/2$, equation (22a) yields $c_1 = 0$, and (22b) becomes

$$(23) \quad (s+1)sc_s + c_{s-2} = 0.$$

From this and $c_1 = 0$ we obtain $c_3 = 0$, $c_5 = 0$, etc. Solving (23) for c_s and setting $s = 2p$, we get

$$c_{2p} = -\frac{c_{2p-2}}{2p(2p+1)} \quad (p=1, 2, \dots).$$

Hence the nonzero coefficients are

$$c_2 = -\frac{c_0}{3!}, \quad c_4 = -\frac{c_2}{4 \cdot 5} = \frac{c_0}{5!}, \quad c_6 = -\frac{c_4}{6!}, \quad \text{etc.,}$$

and we obtain the solution

$$(24) \quad y_1(x) = c_0 x^{\frac{1}{4}} \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{(2m+1)!} = c_0 x^{-\frac{1}{4}} \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m+1}}{(2m+1)!} = c_0 \frac{\sin x}{\sqrt{x}}.$$

From (19) we see that a second independent solution is of the form

$$y_2(x) = ky_1(x) \ln x + x^{-\frac{1}{4}} \sum_{m=0}^{\infty} C_m x^m.$$

If we substitute this and the derivatives into the differential equation, we see that the three expressions involving $\ln x$ and the expressions ky_1 and $-ky_1$ drop out. Simplifying the remaining equation, we thus obtain

$$2kxy_1' + \sum_{m=0}^{\infty} m(m-1)C_m x^{m-\frac{1}{4}} + \sum_{m=0}^{\infty} C_m x^{m+\frac{3}{4}} = 0.$$

From (24) we find $2kxy' = -kc_0 x^{-\frac{1}{4}} + \dots$. Since there is no further term involving $x^{\frac{1}{4}}$ and $c_0 \neq 0$, we must have $k = 0$. The sum of the coefficients of the power $x^{s-\frac{1}{4}}$ is

$$s(s-1)C_s + C_{s-2} \quad (s = 2, 3, \dots).$$

Equating this to zero and solving for C_s , we have

$$C_s = -\frac{C_{s-2}}{s(s-1)} \quad (s = 2, 3, \dots).$$

We thus obtain

$$\begin{aligned} C_2 &= -\frac{C_0}{2!}, & C_4 &= -\frac{C_2}{4 \cdot 3} = \frac{C_0}{4!}, & C_6 &= -\frac{C_4}{6!}, & \text{etc.,} \\ C_3 &= -\frac{C_1}{3!}, & C_5 &= -\frac{C_3}{5 \cdot 4} = \frac{C_1}{5!}, & C_7 &= -\frac{C_5}{7!}, & \text{etc.} \end{aligned}$$

We may take $C_1 = 0$, because the odd powers would yield $C_1 y_1 / c_0$. Then

$$(25) \quad y_2(x) = C_0 x^{-\frac{1}{4}} \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{(2m)!} = C_0 \frac{\cos x}{\sqrt{x}}.$$

$y_1(x)$ and $y_2(x)$ constitute a fundamental system of solutions for all $x \neq 0$.

Example 5 (Cauchy equation). In Sec. 2.7 we have seen that the Cauchy equation

$$(26) \quad x^2 y'' + a_0 x y' + b_0 y = 0 \quad (a_0, b_0 \text{ constant})$$

can be solved by substituting $y = x^r$ into (26). This yields the auxiliary equation

$$r^2 + (a_0 - 1)r + b_0 = 0.$$

Obviously, this is the indicial equation. If $r_1 \neq r_2$, then $y_1 = x^{r_1}$, $y_2 = x^{r_2}$ form a fundamental system; this corresponds to our present Cases 1 and 3. If $r_1 = r_2 = r$, then the solutions

$$y_1 = x^r, \quad y_2 = y_1 \ln x = x^r \ln x$$

form a fundamental system, as follows by substituting (16) into (26), in agreement with the result in Sec. 2.7.

Problems

Find a fundamental system of solutions of the following differential equations. Try to identify the series obtained by the Frobenius method as expansions of known functions.

1. $x^2y'' + 4xy' + (x^2 + 2)y = 0$
2. $2x(1 - x)y'' + (1 + x)y' - y = 0$
3. $y'' + xy' + (1 - 2x^{-2})y = 0$
4. $xy'' + 2y' + 4xy = 0$
5. $xy'' + 3y' + 4x^3y = 0$
6. $xy'' - y = 0$
7. $x^2y'' - 5xy' + 9y = 0$
8. $xy'' + 2y' + xy = 0$
9. $xy'' + y' - xy = 0$
10. $(x - 1)xy'' + (4x - 2)y' + 2y = 0$
11. $x^2y'' + xy' - 4y = 0$
12. $2x^2y'' - xy' - 2y = 0$
13. $x^2y'' + 6xy' + (6 - 4x^2)y = 0$
14. $x^2y'' - 2y = 0$
15. $xy'' + (1 - 2x)y' + (x - 1)y = 0$
16. $xy'' + (2 - 2x)y' + (x - 2)y = 0$

17. (Gauss's hypergeometric equation) The differential equation

$$(27) \quad x(1 - x)y'' + [c - (a + b + 1)x]y' - aby = 0$$

(where a, b, c are constants) is known as *Gauss's hypergeometric equation*. Show that the corresponding indicial equation has the roots $r_1 = 0$ and $r_2 = 1 - c$. Show that for $r_1 = 0$ the Frobenius method yields

$$(28) \quad y_1(x) = 1 + \frac{ab}{1! c} x + \frac{a(a+1)b(b+1)}{2! c(c+1)} x^2 + \frac{a(a+1)(a+2)b(b+1)(b+2)}{3! c(c+1)(c+2)} x^3 + \dots$$

$$(c \neq 0, -1, -2, \dots)$$

This series is called the **hypergeometric series**. Its sum $y_1(x)$ is usually denoted by $F(a, b, c; x)$ and is called the **hypergeometric function**.

18. Prove that the series (28) converges for $|x| < 1$.
19. Show that when $a = b = c = 1$, the series (28) reduces to the geometric series.
20. For what values of a and b does (28) reduce to a polynomial?
21. Show that to $r_2 = 1 - c$ in Prob. 17 there corresponds the solution

$$(29) \quad y_2(x) = x^{1-c} \left(1 + \frac{(a-c+1)(b-c+1)}{1!(-c+2)} x + \frac{(a-c+1)(a-c+2)(b-c+1)(b-c+2)}{2!(-c+2)(-c+3)} x^2 + \dots \right)$$

$$(c \neq 2, 3, 4, \dots)$$

22. Show that in Prob. 21,

$$y_2(x) = x^{1-c} F(a - c + 1, b - c + 1, 2 - c; x).$$

23. Show that if $c \neq 0, \pm 1, \pm 2, \dots$, the functions (28) and (29) constitute a fundamental system of solutions of (27).

Many elementary functions are special cases of $F(a, b, c; x)$. Prove

$$24. \quad \frac{1}{1-x} = F(1, 1, 1; x) = F(1, b, b; x) = F(a, 1, a; x)$$

$$25. \quad (1+x)^n = F(-n, b, b; -x), \quad (1-x)^n = 1 - nx F(1-n, 1, 2; x)$$

26. $\arctan x = xF(\frac{1}{2}, 1, \frac{3}{2}; -x^2)$

27. $\arcsin x = xF(\frac{1}{2}, \frac{1}{2}, \frac{3}{2}; x^2)$

28. $\ln(1+x) = xF(1, 1, 2; -x), \quad \ln \frac{1+x}{1-x} = 2xF(\frac{1}{2}, 1, \frac{3}{2}; x^2)$

29. Consider the differential equation

(30) $(t^2 + At + B)\ddot{y} + (Ct + D)\dot{y} + Ky = 0$

where A, B, C, D, K are constants and $t^2 + At + B$ has distinct zeros t_1 and t_2 . Show that by introducing the new independent variable $x = (t - t_1)/(t_2 - t_1)$ the equation (30) becomes the hypergeometric equation where $Ct_1 + D = -c(t_2 - t_1)$, $C = a + b + 1$, $K = ab$.

Solve, in terms of hypergeometric functions, the following equations.

30. $2x(1-x)y'' + (1-5x)y' - y = 0$

31. $2x(1-x)y'' - (1+5x)y' - y = 0$

32. $5x(1-x)y'' + (4-x)y' + y = 0$

33. $4(t^2 - 3t + 2)\ddot{y} - 2\dot{y} + y = 0$

34. $3t(1+t)\ddot{y} + t\dot{y} - y = 0$

35. $2(t^2 - 5t + 6)\ddot{y} + (2t - 3)\dot{y} - 8y = 0$

3.5 Bessel's Equation. Bessel Functions of the First Kind

One of the most important differential equations in applied mathematics is **Bessel's differential equation**⁸

(1) $x^2y'' + xy' + (x^2 - \nu^2)y = 0$

where the parameter ν is a given number. We assume that ν is real and non-negative. This equation is of the type characterized in Theorem 1, Sec. 3.4, and therefore has a solution of the form

(2) $y(x) = \sum_{m=0}^{\infty} c_m x^{m+r} \quad (c_0 \neq 0).$

Substituting this expression and its first and second derivatives into Bessel's equation, we readily obtain

$$\begin{aligned} \sum_{m=0}^{\infty} (m+r)(m+r-1)c_m x^{m+r} + \sum_{m=0}^{\infty} (m+r)c_m x^{m+r} \\ + \sum_{m=0}^{\infty} c_m x^{m+r+2} - \nu^2 \sum_{m=0}^{\infty} c_m x^{m+r} = 0. \end{aligned}$$

⁸ FRIEDRICH WILHELM BESSEL (1784–1846), German astronomer and mathematician.

By equating the sum of the coefficients of x^{s+r} to zero we find

$$(a) \quad r(r-1)c_0 + rc_0 - \nu^2 c_0 = 0 \quad (s=0)$$

$$(3) \quad (b) \quad (r+1)rc_1 + (r+1)c_1 - \nu^2 c_1 = 0 \quad (s=1)$$

$$(c) \quad (s+r)(s+r-1)c_s + (s+r)c_s + c_{s-2} - \nu^2 c_s = 0 \quad (s=2, 3, \dots).$$

From (3a) we obtain the indicial equation

$$(4) \quad (r+\nu)(r-\nu) = 0.$$

The roots are $r_1 = \nu$ (≥ 0) and $r_2 = -\nu$.

Let us first determine a solution corresponding to the root r_1 . For this value of r , equation (3b) yields $c_1 = 0$. Equation (3c) may be written

$$(s+r+\nu)(s+r-\nu)c_s + c_{s-2} = 0,$$

and for $r = \nu$ this takes the form

$$(5) \quad (s+2\nu)sc_s + c_{s-2} = 0.$$

Since $c_1 = 0$ and $\nu \geq 0$, it follows that $c_3 = 0$, $c_5 = 0$, \dots , successively. If we set $s = 2m$ in (5), we obtain

$$(6) \quad c_{2m} = -\frac{1}{2^m m(\nu+m)} c_{2m-2}, \quad m = 1, 2,$$

and may determine the coefficients c_2, c_4, \dots successively.

c_0 is arbitrary. It is customary to put

$$(7) \quad c_0 = \frac{1}{2^0 \Gamma(\nu+1)}$$

where $\Gamma(\nu+1)$ is the **gamma function**. The properties of this important function will be considered in Sec. 17.1. For the present purpose it suffices to know that $\Gamma(\alpha)$ is defined by the integral

$$(8) \quad \Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1} dt \quad (\alpha > 0).$$

By integration by parts we obtain

$$\Gamma(\alpha+1) = \int_0^\infty e^{-t} t^\alpha dt = -e^{-t} t^\alpha \Big|_0^\infty + \alpha \int_0^\infty e^{-t} t^{\alpha-1} dt.$$

The first expression on the right is zero, and the integral on the right is $\Gamma(\alpha)$. This yields the basic relation

$$(9) \quad \Gamma(\alpha+1) = \alpha \Gamma(\alpha).$$

Since

$$\Gamma(1) = \int_0^\infty e^{-t} dt = 1$$

we conclude from (9) that

$$\Gamma(2) = \Gamma(1) = 1!, \quad \Gamma(3) = 2\Gamma(2) = 2!, \dots$$

and in general

$$(10) \quad \Gamma(k+1) = k! \quad (k = 0, 1, \dots).$$

This shows that the gamma function may be regarded as a generalization of the factorial function known from elementary calculus.

Returning to our problem, we obtain from (6), (7), and (9)

$$c_2 = -\frac{c_0}{2^2(\nu+1)} = -\frac{1}{2^{2+\nu}1!\Gamma(\nu+2)}$$

$$c_4 = -\frac{c_2}{2^42(\nu+2)} = \frac{1}{2^{4+\nu}2!\Gamma(\nu+3)}$$

and so on; in general,

$$(11) \quad c_{2m} = \frac{(-1)^m}{2^{2m+\nu}m!\Gamma(\nu+m+1)}.$$

By inserting these coefficients in (2) and remembering that $c_1 = 0, c_3 = 0, \dots$, we obtain a particular solution of Bessel's equation which is denoted by $J_\nu(x)$:

$$(12) \quad J_\nu(x) = x^\nu \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m+\nu} m! \Gamma(\nu+m+1)} = x^\nu \sum_{m=0}^{\infty} \frac{(-1)^m}{2^{2m} m! \Gamma(\nu+m+1)} \left(\frac{x}{2}\right)^{2m}$$

This solution of (1) is known as the **Bessel function of the first kind of order ν** . The ratio test shows that the series converges for all x .

Integer values of ν are frequently denoted by n . Thus, for $n \geq 0$,

$$(13) \quad J_n(x) = x^n \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m+n} m! (n+m)!}.$$

Replacing ν by $-\nu$ in (12) we have

$$(14) \quad J_{-\nu}(x) = x^{-\nu} \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m-\nu} m! \Gamma(m-\nu+1)}.$$

Since Bessel's equation involves x^2 , the functions J_ν and $J_{-\nu}$ are solutions of the equation for the same ν . If ν is not an integer, they are linearly independent, because the first term in (12) and the first term in (14) are finite nonzero multiples of x^ν and $x^{-\nu}$, respectively. This yields the following result.

Theorem 1 (General solution). *If ν is not an integer, a general solution of Bessel's equation for all $x \neq 0$ is*

$$(15) \quad y(x) = a_1 J_\nu(x) + a_2 J_{-\nu}(x).$$

However, if ν is an integer, then (15) is not a general solution. This follows from *

Table 14. Bessel Functions $J_0(x)$ and $J_1(x)$.(Zeros of J_0 see in Sec. 9.10. For more extensive tables see Ref. [A7] in Appendix 1.)

x	$J_0(x)$	$J_1(x)$	x	$J_0(x)$	$J_1(x)$	x	$J_0(x)$	$J_1(x)$
0.0	1.0000	0.0000	3.0	-0.2601	0.3391	6.0	0.1506	-0.2767
0.1	0.9975	0.0499	3.1	-0.2921	0.3009	6.1	0.1773	-0.2559
0.2	0.9900	0.0995	3.2	-0.3202	0.2613	6.2	0.2017	-0.2329
0.3	0.9776	0.1483	3.3	-0.3443	0.2207	6.3	0.2238	-0.2081
0.4	0.9604	0.1960	3.4	-0.3643	0.1792	6.4	0.2433	-0.1816
0.5	0.9385	0.2423	3.5	-0.3801	0.1374	6.5	0.2601	-0.1538
0.6	0.9120	0.2867	3.6	-0.3918	0.0955	6.6	0.2740	-0.1250
0.7	0.8812	0.3290	3.7	-0.3992	0.0538	6.7	0.2851	-0.0953
0.8	0.8463	0.3688	3.8	-0.4026	0.0128	6.8	0.2931	-0.0652
0.9	0.8075	0.4059	3.9	-0.4018	-0.0272	6.9	0.2981	-0.0349
1.0	0.7652	0.4401	4.0	-0.3971	-0.0660	7.0	0.3001	-0.0047
1.1	0.7196	0.4709	4.1	-0.3887	-0.1033	7.1	0.2991	0.0252
1.2	0.6711	0.4983	4.2	-0.3766	-0.1386	7.2	0.2951	0.0543
1.3	0.6201	0.5220	4.3	-0.3610	-0.1719	7.3	0.2882	0.0826
1.4	0.5669	0.5419	4.4	-0.3423	-0.2028	7.4	0.2786	0.1096
1.5	0.5118	0.5579	4.5	-0.3205	-0.2311	7.5	0.2663	0.1352
1.6	0.4554	0.5699	4.6	-0.2961	-0.2566	7.6	0.2516	0.1592
1.7	0.3980	0.5778	4.7	-0.2693	-0.2791	7.7	0.2346	0.1813
1.8	0.3400	0.5815	4.8	-0.2404	-0.2985	7.8	0.2154	0.2014
1.9	0.2818	0.5812	4.9	-0.2097	-0.3147	7.9	0.1944	0.2192
2.0	0.2239	0.5767	5.0	-0.1776	-0.3276	8.0	0.1717	0.2346
2.1	0.1666	0.5683	5.1	-0.1443	-0.3371	8.1	0.1475	0.2476
2.2	0.1104	0.5560	5.2	-0.1103	-0.3432	8.2	0.1222	0.2580
2.3	0.0555	0.5399	5.3	-0.0758	-0.3460	8.3	0.0960	0.2657
2.4	0.0025	0.5202	5.4	-0.0412	-0.3453	8.4	0.0692	0.2708
2.5	-0.0484	0.4971	5.5	-0.0068	-0.3414	8.5	0.0419	0.2731
2.6	-0.0968	0.4708	5.6	0.0270	-0.3343	8.6	0.0146	0.2728
2.7	-0.1424	0.4416	5.7	0.0599	-0.3241	8.7	-0.0125	0.2697
2.8	-0.1850	0.4097	5.8	0.0917	-0.3110	8.8	-0.0392	0.2641
2.9	-0.2243	0.3754	5.9	0.1220	-0.2951	8.9	-0.0653	0.2559

 $J_1(x) = 0$ for $x = 0, 3.832, 7.016, 10.173, 13.324, \dots$

Theorem 2 (Linear dependence). For integer $\nu = r$, the Bessel functions $J_n(x)$ and $J_{-n}(x)$ are linearly dependent, because

$$(16) \quad J_{-n}(x) = (-1)^n J_n(x) \quad (n = 1, 2, \dots).$$

Proof. We use (14) and let ν approach a positive integer n . Then the gamma functions in the coefficients of the first n terms become infinite (cf. Fig. 373 in Sec. 17.1), the coefficients become zero, and the summation starts with $m = n$. Since in this case $\Gamma(m - n + 1) = (m - n)!$ [cf. (10)], we obtain

$$J_{-n}(x) = \sum_{m=n}^{\infty} \frac{(-1)^m x^{2m-n}}{2^{2m-n} m! (m-n)!} = \sum_{s=0}^{\infty} \frac{(-1)^{n+s} x^{2s+n}}{2^{2s+n} (n+s)! s!}$$

where $m = n + s$ and $s = m - n$. From (13) we see that the last series represents $(-1)^n J_n(x)$. This completes the proof.

A general solution of the Bessel equation with integer $\nu = n$ will be obtained in Sec. 3.7.

Problems

1. Show that (12) may also be written in the form

$$(12') \quad J_\nu(x) = \left(\frac{x}{2}\right)^\nu \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(\nu + m + 1)} \left(\frac{x}{2}\right)^{2m}.$$

2. Prove that the series in (12) converges for all x .

3. Using (13), show that $J_n(x)$ is even for even n and odd for odd n .

4. Show that (cf. Fig. 80)

$$(17) \quad J_0(x) = 1 - \frac{x^2}{2^2(1!)^2} + \frac{x^4}{2^4(2!)^2} - \frac{x^6}{2^6(3!)^2} + \dots$$

5. Show that (cf. Fig. 80)

$$(18) \quad J_1(x) = \frac{x}{2} - \frac{x^3}{2^3 1! 2!} + \frac{x^5}{2^5 2! 3!} - \frac{x^7}{2^7 3! 4!} + \dots$$

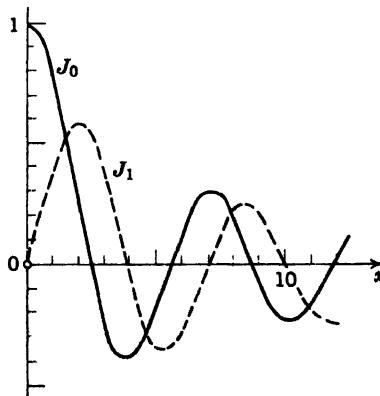


Fig. 80. Bessel functions of the first kind.

6. Obtain (14) directly by setting $r = -\nu$ in (3a) and (3b).
7. Obtain $J_{-1}(x)$ from (14) and compare it with (18).
8. Show that $J_0'(x) = -J_1(x)$.
9. (Elimination of first derivative) Consider the differential equation

$$y'' + f(x)y' + g(x)y = 0.$$

Substitute $y(x) = u(x)v(x)$ and determine v such that the resulting second-order differential equation does not involve u' .

10. Show that for the Bessel equation the substitution in Prob. 9 is $y = ux^{-\frac{1}{2}}$ and yields

$$(19) \quad x^2u'' + (x^2 + \frac{1}{4} - \nu^2)u = 0.$$

11. Using (19), find a general solution of (1) with $\nu = 1/2$.
12. Using (12) and the relation $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ (which will be proved in Sec. 17.1), show that

$$(20) \quad J_{\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}} \sin x, \quad J_{-\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}} \cos x.$$

The following problems illustrate that various differential equations can be reduced to Bessel's equation. Using the indicated substitutions, solve, in terms of Bessel functions, the following equations.

13. $x^2y'' + xy' + (x^2 - \frac{1}{9})y = 0$
14. $x^2y'' + xy' + \frac{1}{4}(x^2 - 1)y = 0 \quad (x = 2z)$
15. $x^2y'' + xy' + (\lambda^2x^2 - \nu^2)y = 0 \quad (x = z/\lambda)$
16. $x^2y'' + xy' + \frac{1}{4}(x - \nu^2)y = 0 \quad (\sqrt{x} = z)$
17. $x^2y'' + xy' + 4(x^4 - \nu^2)y = 0 \quad (x^2 = z)$
18. $xy'' + (1 + 2\nu)y' + xy = 0 \quad (y = x^{-\nu}u)$
19. $y'' + xy = 0 \quad (y = x^{\frac{1}{2}\nu}u, \frac{2}{3}x^{3/2} = z)$
20. $x^2y'' + (1 - 2\nu)xy' + \nu^2(x^{2\nu} + 1 - \nu^2)y = 0 \quad (y = x^\nu u, x^\nu = z)$

21. Show that for small $|x|$ we have $J_0(x) \approx 1 - \frac{1}{4}x^2$. Using this formula, compute $J_0(x)$ for $x = 0.1, 0.2, \dots, 1.0$ and determine the relative error by comparing with Table 14.
22. How many terms of (17) are needed to compute $J_0(1)$ with an error less than 1 unit of the fifth decimal place? How many terms would be needed to compute $\ln 2$ from the Maclaurin series of $\ln(1+x)$ with the same degree of accuracy?
23. Using (17), show that the smallest positive root of $J_0(x) = 0$ lies between 2 and $\sqrt{8}$ (≈ 2.83). (The root is $x = 2.404 \dots$)
24. It can be shown that for large x ,

$$(21) \quad \begin{aligned} J_{2n}(x) &\approx (-1)^n(\pi x)^{-\frac{1}{2}} (\cos x + \sin x) \\ J_{2n+1}(x) &\approx (-1)^{n+1}(\pi x)^{-\frac{1}{2}} (\cos x - \sin x). \end{aligned}$$

Using (21), graph $J_0(x)$ for large x , compute approximate values of the first five positive zeros of $J_0(x)$, and compare them with the more accurate values 2.405, 5.520, 8.654, 11.792, 14.931.

25. Do the same task as in Prob. 24, for $J_1(x)$.

3.6 Further Properties of Bessel Functions of the First Kind

Bessel functions play an important role in various engineering problems. Some typical applications will be considered in connection with partial differential equations. In working with Bessel functions it is necessary to know that these functions satisfy various relations, and, in fact, the number of formulas for Bessel functions is incredibly large; the student may look into Ref. [B9], the standard book on Bessel functions, or Ref. [A5] in Appendix 1.

To show how to obtain properties of functions from their series representations we shall now consider some important relations for Bessel functions of the first kind.

From (12) in Sec. 3.5 we obtain

$$x^v J_v(x) = \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m+2v}}{2^{2m+v} m! \Gamma(v+m+1)}.$$

By differentiating and then using (9), Sec. 3.5, we readily find

$$(x^v J_v)' = \sum_{m=0}^{\infty} \frac{(-1)^m 2(m+v)x^{2m+2v-1}}{2^{2m+v} m! \Gamma(v+m+1)} = x^v x^{v-1} \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m+v-1} m! \Gamma(v+m)}.$$

From (12), Sec. 3.5, it follows that the last expression is $x^v J_{v-1}(x)$ and, therefore,

$$(1) \quad \frac{d}{dx} [x^v J_v(x)] = x^v J_{v-1}(x).$$

Similarly, from (12) in Sec. 3.5 it follows that

$$(x^{-v} J_v)' = \sum_{m=1}^{\infty} \frac{(-1)^m x^{2m-1}}{2^{2m+v-1} (m-1)! \Gamma(v+m+1)} = \sum_{s=0}^{\infty} \frac{(-1)^{s+1} x^{2s+1}}{2^{2s+v+1} s! \Gamma(v+s+2)}$$

(where $m = s + 1$). Since the expression on the right is $-x^{-v} J_{v+1}(x)$, this proves

$$(2) \quad \frac{d}{dx} [x^{-v} J_v(x)] = -x^{-v} J_{v+1}(x).$$

Writing these formulas at length and multiplying the last by x^{2v} , we get

$$(1') \quad vx^{v-1} J_v + x^v J_v' = x^v J_{v-1}$$

$$(2') \quad -vx^{v-1} J_v + x^v J_v' = -x^v J_{v+1}.$$

Subtracting (2') from (1') and dividing the result by x^v , we find

$$(3) \quad J_{v-1}(x) + J_{v+1}(x) = \frac{2v}{x} J_v(x).$$

Adding (1') and (2') and dividing the result by x^v , we obtain

$$(4) \quad J_{v-1}(x) - J_{v+1}(x) = 2J_v'(x).$$

The recurrence relations (3) and (4) are of great practical importance. For example, (3) serves to express Bessel functions of high orders in terms of functions of low orders, and this is useful for computing tables.

Example 1. Express $J_3(x)$ in terms of $J_0(x)$ and $J_1(x)$. From (3) we obtain

$$\begin{aligned} J_2(x) &= \frac{2}{x} J_1(x) - J_0(x), \\ J_3(x) &= \frac{4}{x} J_2(x) - J_1(x) = \left(\frac{8}{x^2} - 1 \right) J_1(x) - \frac{4}{x} J_0(x). \end{aligned}$$

Example 2 (Elementary functions). From Prob. 12 in the last section we know that

$$(5) \quad (a) \quad J_{\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}} \sin x \quad (b) \quad J_{-\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}} \cos x$$

(cf. also Ex. 4 in Sec. 3.4). By applying our recurrence relation (3) we thus obtain

$$(6) \quad \begin{aligned} J_{\frac{3}{2}}(x) &= \frac{1}{x} J_{\frac{1}{2}}(x) - J_{-\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}} \left(\frac{\sin x}{x} - \cos x \right), \\ J_{-\frac{3}{2}}(x) &= -\frac{1}{x} J_{-\frac{1}{2}}(x) - J_{\frac{1}{2}}(x) = -\sqrt{\frac{2}{\pi x}} \left(\frac{\cos x}{x} + \sin x \right), \end{aligned}$$

and so on. Continuing in this fashion we see that for $v = \pm \frac{1}{2}, \pm \frac{3}{2}, \pm \frac{5}{2}, \dots$ the Bessel functions $J_v(x)$ are elementary functions.

Problems

1. Derive Bessel's equation from (1) and (2).
2. Carry out the details of the derivation of (3) and (4) from (1) and (2).
3. Derive (5b) from (5a) and (1).
4. Using (3), prove (16) in Sec. 3.5 by induction.
5. Using (1) and (2), obtain the right-hand sides of (6) from (5).

Express in terms of $J_0(x)$ and $J_1(x)$:

6. $J_2(x)$

7. $J_3(x)$

8. $J_4(x)$

9. Using the result of Prob. 6, compute $J_2(x)$ and $J_{-2}(x)$ for $x = 0, 0.1, 0.2, \dots, 1.0$ from the table in Sec. 3.5.

Using (1) to (4), show that

10. $J_0(x) = -J_1(x)$

11. $J_n'' = \frac{1}{4}(J_{n-2} - 2J_n + J_{n+2})$

12. $J_1'(x) = J_0(x) - \frac{1}{x} J_1(x)$

13. $J_2'(x) = \left(1 - \frac{4}{x^2} \right) J_1(x) + \frac{2}{x} J_0(x)$

14. $\int x^v J_{v-1}(x) dx = x^v J_v(x) + c, \quad \int x^{-v} J_{v+1}(x) dx = -x^{-v} J_v(x) + c$

15. $\int J_{v+1}(x) dx = \int J_{v-1}(x) dx - 2J_v(x)$

Using the formulas in Probs. 14 and 15 (and, if necessary, integration by parts), evaluate

16. $\int x J_0(x) dx$

17. $\int J_1(x) dx$

18. $\int J_3(x) dx$

19. $\int J_5(x) dx$

20. $\int x^{-3} J_4(x) dx$

21. $\int x^3 J_0(x) dx$

Show that

22. $\int J_2(x) dx = \int J_0(x) dx - 2 J_1(x)$

23. $\int x J_1(x) dx = \int J_0(x) dx - x J_0(x)$

(The remaining integral, which cannot be evaluated in finite form, has been tabulated; cf. Ref. [B9] in Appendix 1.)

24. Using (1) and (2) and Rolle's theorem, show that between two consecutive zeros of $J'_0(x)$ there is precisely one zero of $J_1(x)$.

25. Show that between any two consecutive positive zeros of $J_n(x)$ there is precisely one zero of $J_{n+1}(x)$.

3.7 Bessel Functions of the Second Kind

For integer $v = n$, the Bessel functions $J_n(x)$ and $J_{-n}(x)$ are linearly dependent (Sec. 3.5) and do not form a fundamental system. We shall now obtain a second independent solution, starting with the case $n = 0$. In this case Bessel's equation may be written

$$(1) \quad xy'' + y' + xy = 0,$$

the indicial equation has the double root $r = 0$, and we see from (16) in Sec. 3.4, that the desired solution must have the form

$$(2) \quad y_2(x) = J_0(x) \ln x + \sum_{m=1}^{\infty} A_m x^m.$$

When we substitute y_2 and its derivatives

$$y_2' = J_0' \ln x + \frac{J_0}{x} + \sum_{m=1}^{\infty} m A_m x^{m-1}$$

$$y_2'' = J_0'' \ln x + \frac{2J_0'}{x} - \frac{J_0}{x^2} + \sum_{m=1}^{\infty} m(m-1) A_m x^{m-2}$$

into (1), the logarithmic terms disappear because J_0 is a solution of (1), the other two terms containing J_0 cancel, and we find

$$2J_0' + \sum_{m=1}^{\infty} m(m-1) A_m x^{m-1} + \sum_{m=1}^{\infty} m A_m x^{m-1} + \sum_{m=1}^{\infty} A_m x^{m+1} = 0.$$

From (12) in Sec. 3.5 we obtain the power series of J_0' in the form

$$J_0'(x) = \sum_{m=1}^{\infty} \frac{(-1)^m 2mx^{2m-1}}{2^{2m}(m!)^2} = \sum_{m=1}^{\infty} \frac{(-1)^m x^{2m-1}}{2^{2m-1} m! (m-1)!}.$$

By inserting this series we have

$$\sum_{m=1}^{\infty} \frac{(-1)^m x^{2m-1}}{2^{2m-2} m! (m-1)!} + \sum_{m=1}^{\infty} m^2 A_m x^{m-1} + \sum_{m=1}^{\infty} A_m x^{m+1} = 0.$$

We first show that the A_m with odd subscripts are all zero. The coefficient of the power x^0 is A_1 , and so, $A_1 = 0$. By equating the sum of the coefficients of the power x^{2s} to zero we obtain

$$(2s+1)^2 A_{2s+1} + A_{2s-1} = 0, \quad s = 1, 2, \dots$$

Since $A_1 = 0$, we thus obtain $A_3 = 0$, $A_5 = 0, \dots$, successively. Equating the sum of the coefficients of x^{2s+1} to zero, we have

$$-1 + 4A_2 = 0 \quad (s = 0)$$

$$\frac{(-1)^{s+1}}{2^{2s}(s+1)! s!} + (2s+2)^2 A_{2s+2} + A_{2s} = 0 \quad (s = 1, 2, \dots)$$

For $s = 0$ this yields $4A_2 = 1$ or $A_2 = \frac{1}{4}$. For $s = 1$ we find

$$\frac{1}{8} + 16A_4 + A_2 = 0 \quad \text{or} \quad A_4 = -\frac{1}{128}$$

and in general

$$(3) \quad A_{2m} = \frac{(-1)^{m-1}}{2^{2m}(m!)^2} \left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{m} \right), \quad m = 1, 2, \dots$$

Using the short notation

$$(4) \quad h_m = 1 + \frac{1}{2} + \dots + \frac{1}{m}$$

and inserting (3) and $A_1 = A_3 = \dots = 0$ into (2), we obtain the result

$$(5) \quad \begin{aligned} y_2(x) &= J_0(x) \ln x + \sum_{m=1}^{\infty} \frac{(-1)^{m-1} h_m}{2^{2m}(m!)^2} x^{2m} \\ &= J_0(x) \ln x + \frac{1}{4} x^2 - \frac{1}{128} x^4 + \dots \end{aligned}$$

Since J_0 and y_2 are linearly independent functions, they form a fundamental system of (1). Of course, another fundamental system is obtained by replacing y_2 by an independent particular solution of the form $a(y_2 + bJ_0)$ where a ($\neq 0$) and b are constants. It is customary to choose $a = 2/\pi$ and $b = \gamma - \ln 2$ where $\gamma = 0.5772 \dots$ is the so-called *Euler constant*, which is defined as the limit of

$$1 + \frac{1}{2} + \dots + \frac{1}{s} - \ln s$$

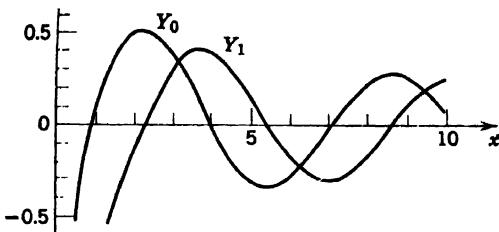


Fig. 81. Bessel functions of the second kind.

as s approaches infinity. The standard particular solution thus obtained is known as the **Bessel function of the second kind of order zero** (Fig. 81) or **Neumann's function of order zero** and is denoted by $Y_0(x)$. Thus

$$(6) \quad Y_0(x) = \frac{2}{\pi} \left[J_0(x) \left(\ln \frac{x}{2} + \gamma \right) + \sum_{m=1}^{\infty} \frac{(-1)^{m-1} h_m}{2^{2m} (m!)^2} x^{2m} \right],$$

where h_m is defined by (4).

If $\nu = n = 1, 2, \dots$, a second solution can be obtained by similar manipulations, starting from (19), Sec. 3.4. It turns out that also in these cases the solution contains a logarithmic term.

The situation is not yet completely satisfactory, because the second solution is defined differently, depending on whether the order ν is integral or not. To provide uniformity of formalism and numerical tabulation, it is desirable to adopt a form of the second solution which is valid for all values of the order. This is the reason for introducing a standard second solution $Y_\nu(x)$ defined for all ν by the formula

$$(7) \quad (a) \quad Y_\nu(x) = \frac{1}{\sin \nu\pi} [J_\nu(x) \cos \nu\pi - J_{-\nu}(x)]$$

$$(b) \quad Y_n(x) = \lim_{\nu \rightarrow n} Y_\nu(x).$$

This function is known as the **Bessel function of the second kind of order ν** or **Neumann's function⁹ of order ν** .

For nonintegral order ν , the function $Y_\nu(x)$ is evidently a solution of Bessel's equation because $J_\nu(x)$ and $J_{-\nu}(x)$ are solutions of that equation. Since for those ν the solutions J_ν and $J_{-\nu}$ are linearly independent and Y_ν involves $J_{-\nu}$, the functions J_ν and Y_ν are linearly independent. Furthermore, it can be shown that the limit in (7b) exists and Y_n is a solution of Bessel's equation for integral order; cf. Ref. [B9] in Appendix 1. We shall see that the series development of $Y_n(x)$ contains a logarithmic term. Hence $J_n(x)$

⁹ CARL NEUMANN (1832–1925), German mathematician and physicist. The solutions $Y_\nu(x)$ are sometimes denoted by $N_\nu(x)$; in Ref. [B9] they are called Weber's functions. Euler's constant in (6) is often denoted by C or in γ (instead of γ).

and $Y_n(x)$ are linearly independent solutions of Bessel's equation. The series development of $Y_n(x)$ can be obtained by inserting the series (12) and (14), Sec. 3.5, for $J_v(x)$ and $J_{-v}(x)$ into (7a) and then letting v approach n ; for details see Ref. [B9]; the result is

$$(8) \quad Y_n(x) = \frac{2}{\pi} J_n(x) \left(\ln \frac{x}{2} + \gamma \right) + \frac{x^n}{\pi} \sum_{m=0}^{\infty} \frac{(-1)^{m-1} (h_m + h_{m+n})}{2^{2m+n} m! (m+n)!} x^{2m} - \frac{x^{-n}}{\pi} \sum_{m=0}^{n-1} \frac{(n-m-1)!}{2^{2m-n} m!} x^{2m} \quad (x > 0, n = 0, 1, \dots),$$

where

$$h_0 = 0, \quad h_s = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{s} \quad (s = 1, 2, \dots)$$

and when $n = 0$ the last sum in (8) is to be replaced by 0. For $n = 0$ the representation (8) takes the form (6). Furthermore, it can be shown that

$$Y_{-n}(x) = (-1)^n Y_n(x).$$

We may formulate our result, as follows.

Theorem 1 (General solution). *A general solution of Bessel's equation for all values of v is*

$$(9) \quad y(x) = C_1 J_v(x) + C_2 Y_v(x).$$

We finally mention that there is a practical need for solutions of Bessel's equation which are complex for real values of x . For this reason the solutions

$$(10) \quad \begin{aligned} H_v^{(1)}(x) &= J_v(x) + i Y_v(x) \\ H_v^{(2)}(x) &= J_v(x) - i Y_v(x) \end{aligned}$$

are frequently used and have been tabulated (cf. Ref [A6] in Appendix 1). These linearly independent functions are called **Bessel functions of the third kind of order v** or **first and second Hankel functions¹⁰ of order v** .

Table 15. Bessel Functions $Y_0(x)$ and $Y_1(x)$.

(For more extensive tables see Ref. [A7].)

x	$Y_0(x)$	$Y_1(x)$	x	$Y_0(x)$	$Y_1(x)$	x	$Y_0(x)$	$Y_1(x)$
0.0	($-\infty$)	($-\infty$)	2.5	0.498	0.146	5.0	-0.309	0.148
0.5	-0.445	-1.471	3.0	0.377	0.325	5.5	-0.340	-0.024
1.0	0.088	-0.781	3.5	0.189	0.410	6.0	-0.288	-0.175
1.5	0.382	-0.412	4.0	-0.017	0.398	6.5	-0.173	-0.274
2.0	0.510	-0.107	4.5	-0.195	0.301	7.0	-0.026	-0.303

¹⁰ HERMANN HANKEL (1839–1873), German mathematician.

Problems

Using the indicated substitutions, find a general solution.

1. $xy'' + y' + xy = 0$
2. $x^2y'' + xy' + (x^2 - 1)y = 0$
3. $4xy'' + 4y' + y = 0 \quad (\sqrt{x} = z)$
4. $xy'' - y' + xy = 0 \quad (y = xu)$
5. $xy'' + 3y' + xy = 0 \quad (y = u/x)$
6. $y'' + e^{2x}y = 0 \quad (e^x = z)$
7. $xy'' + 5y' + xy = 0 \quad (y = u/x^2)$
8. $x^2y'' + xy' + 4x^4y = 0 \quad (x^2 = z)$
9. $x^2y'' + xy' + (\lambda^2x^2 - 4)y = 0 \quad (x = z/\lambda)$
10. $x^2y'' - 3xy' + 4(x^4 - 3)y = 0 \quad (y = x^2u, x^2 = z)$
11. $4x^2y'' + (x + 1)y = 0 \quad (y = u\sqrt{x}, \sqrt{x} = z)$
12. $x^2y'' - 5xy' + 9(x^6 - 8)y = 0 \quad (y = x^3u, x^3 = z)$

13. Show that for small $x > 0$ we have $Y_0(x) \approx 2\pi^{-1}(\ln \frac{1}{2}x + \gamma)$.

14. It can be shown that, for large x ,

$$(11) \quad {}^t Y_n(x) \approx \sqrt{2/\pi x} \sin(x - \frac{1}{2}n\pi - \frac{1}{4}\pi).$$

Using (11) and Table 15, graph $Y_0(x)$ and $Y_1(x)$ for $0 < x \leq 15$. Using (11), compute approximate values of the first three positive zeros of $Y_0(x)$ and compare these values with the more accurate values 0.89, 3.96, and 7.09.

15. (Modified Bessel functions) Show that the function $I_v(x) = i^{-v}J_v(ix)$, where $i = \sqrt{-1}$, has the Maclaurin series

$$I_v(x) = i^{-v}J_v(ix) = \sum_{m=0}^{\infty} \frac{x^{2m+v}}{2^{2m+v} m! \Gamma(m+v+1)}.$$

$I_v(x)$ is called the *modified Bessel function of the first kind of order v*.

16. Show that $y = I_v(x)$ is a solution of the differential equation

$$x^2y'' + xy' - (x^2 + v^2)y = 0.$$

17. Show that another solution of the equation in Prob. 16 is

$$K_v(x) = \frac{\pi}{2 \sin v\pi} [I_{-v}(x) - I_v(x)].$$

$K_v(x)$ is called the *modified Bessel function of the third kind* (sometimes: *of the second kind*).

18. Show that $I_v(x)$ is real for all real x (and real v) and is different from zero for all real $x \neq 0$.
19. Show that $I_{-n}(x) = I_n(x)$ where n is any integer.
20. Show that the Hankel functions (10) constitute a fundamental system of solutions for any v .

Chapter 4

Laplace Transformation

The Laplace transformation is a powerful method for solving linear differential equations arising in engineering mathematics. It consists essentially of three steps. In the first step the given differential equation is transformed into an algebraic equation (*subsidiary equation*). Then the latter is solved by purely algebraic manipulations. Finally, the solution of the subsidiary equation is transformed back in such a way that it becomes the required solution of the original differential equation.

In this way the Laplace transformation reduces the problem of solving a differential equation to an algebraic problem. Another advantage is that it takes care of initial conditions without the necessity of first determining the general solution and then obtaining from it a particular solution. Also, when applying the classical method to a nonhomogeneous equation, we must first solve the corresponding homogeneous equation, while the Laplace transformation immediately yields the solution of the nonhomogeneous equation.

In the present chapter we shall consider the Laplace transformation from a practical point of view and illustrate its use by important engineering applications.

Prerequisite for this chapter: Chap. 2.

References: Appendix 1, Part B.

Answers to problems: Appendix 2.

4.1 Laplace Transform. Inverse Transform. Linearity

Let $f(t)$ be a given function which is defined for all positive values of t . We multiply $f(t)$ by e^{-st} and integrate with respect to t from zero to infinity. Then, if the resulting integral exists, it is a function of s , say, $F(s)$.

$$F(s) = \int_0^{\infty} e^{-st} f(t) dt.$$

The function $F(s)$ is called the **Laplace transform**¹ of the original function $f(t)$, and will be denoted by $\mathcal{L}(f)$. Thus

$$(1) \quad F(s) = \mathcal{L}(f) = \int_0^\infty e^{-st} f(t) dt.$$

The described operation on $f(t)$ is called the *Laplace transformation*.

Furthermore the original function $f(t)$ in (1) is called the *inverse transform* or *inverse* of $F(s)$ and will be denoted by $\mathcal{L}^{-1}(F)$; that is, we shall write

$$f(t) = \mathcal{L}^{-1}(F).$$

We shall denote the original function by a lower case letter and its transform by the same letter in capital.

Example 1. Let $f(t) = 1$ when $t > 0$. Then

$$\mathcal{L}(f) = \mathcal{L}(1) = \int_0^\infty e^{-st} dt = -\frac{1}{s} e^{-st} \Big|_0^\infty;$$

hence, when $s > 0$,

$$\mathcal{L}(1) = \frac{1}{s}.$$

Example 2. Let $f(t) = e^{at}$ when $t > 0$, where a is a constant. Then

$$\mathcal{L}(e^{at}) = \int_0^\infty e^{-st} e^{at} dt = \frac{1}{a-s} e^{-(s-a)t} \Big|_0^\infty;$$

consequently, when $s - a > 0$,

$$\mathcal{L}(e^{at}) = \frac{1}{s-a}.$$

We shall see that the Laplace transformation possesses various general properties, in consequence of which transforms of many functions can be obtained in very simple ways. The most important property of the Laplace transformation is stated in the following theorem.

Theorem 1 (Linearity property). *The Laplace transformation is a linear operation, that is, for any functions $f(t)$ and $g(t)$ whose Laplace transforms exist and any constants a and b we have*

$$\mathcal{L}\{af(t) + bg(t)\} = a\mathcal{L}(f) + b\mathcal{L}(g).$$

Proof. By definition,

$$\begin{aligned} \mathcal{L}\{af(t) + bg(t)\} &= \int_0^\infty e^{-st}[af(t) + bg(t)] dt \\ &= a \int_0^\infty e^{-st}f(t) dt + b \int_0^\infty e^{-st}g(t) dt = a\mathcal{L}(f) + b\mathcal{L}(g). \end{aligned}$$

¹PIERRE SIMON DE LAPLACE (1749–1827), great French mathematician, who developed the foundation of potential theory and made important contributions to celestial mechanics and probability theory.

Example 3. Let $f(t) = \cosh at = (e^{at} + e^{-at})/2$. Then, using Theorem 1 and the result of Ex. 2, we find

$$\mathcal{L}(\cosh at) = \frac{1}{2}\mathcal{L}(e^{at}) + \frac{1}{2}\mathcal{L}(e^{-at}) = \frac{1}{2}\left(\frac{1}{s-a} + \frac{1}{s+a}\right);$$

that is, when $s > a$,

$$\mathcal{L}(\cosh at) = \frac{s}{s^2 - a^2}.$$

A short list of some important elementary functions and their Laplace transforms is given in Table 16, and a more extensive list in Table 17, Sec. 4.10. Cf. also the references in Appendix 1, Part B.

Table 16. Some Elementary Functions $f(t)$ and Their Laplace Transforms $\mathcal{L}(f)$.

	$f(t)$	$\mathcal{L}(f)$		$f(t)$	$\mathcal{L}(f)$
1	1	$1/s$	6	e^{at}	$\frac{1}{s-a}$
2	t	$1/s^2$	7	$\cos \omega t$	$\frac{s}{s^2 + \omega^2}$
3	t^2	$2!/s^3$	8	$\sin \omega t$	$\frac{\omega}{s^2 + \omega^2}$
4	t^n ($n = 1, 2, \dots$)	$\frac{n!}{s^{n+1}}$	9	$\cosh at$	$\frac{s}{s^2 - a^2}$
5	t^a (a positive)	$\frac{\Gamma(a+1)}{s^{a+1}}$	10	$\sinh at$	$\frac{a}{s^2 - a^2}$

Once we know the transforms in Table 16, nearly all the transforms we shall need can be obtained through the use of some simple general theorems which we shall consider in the subsequent sections.

Formulas 1, 2, and 3 in Table 16 are special cases of formula 4. Formula 4 follows from formula 5 and $\Gamma(n+1) = n!$, where n is a non-negative integer (cf. (10) in Sec. 3.5). Formula 5 can be proved by starting from the definition

$$\mathcal{L}(t^a) = \int_0^\infty e^{-st} t^a dt,$$

setting $st = x$, and using (8) in Sec. 3.5; then

$$\mathcal{L}(t^a) = \int_0^\infty e^{-x} \left(\frac{x}{s}\right)^a \frac{dx}{s} = \frac{1}{s^{a+1}} \int_0^\infty e^{-x} x^a dx = \frac{\Gamma(a+1)}{s^{a+1}} \quad (s > 0).$$

Formula 6 was proved in Ex. 2. To prove the formulas 7 and 8 we set $a = i\omega$ in formula 6. Then

$$\mathcal{L}(e^{i\omega t}) : \frac{1}{s - i\omega} = \frac{s + i\omega}{(s - i\omega)(s + i\omega)} = \frac{s + i\omega}{s^2 + \omega^2} = \frac{s}{s^2 + \omega^2} + i \frac{\omega}{s^2 + \omega^2}.$$

On the other hand, by Theorem 1,

$$\mathcal{L}(e^{i\omega t}) = \mathcal{L}(\cos \omega t + i \sin \omega t) = \mathcal{L}(\cos \omega t) + i\mathcal{L}(\sin \omega t).$$

Equating the real and imaginary parts of these two equations, we obtain the formulas 7 and 8. Formula 9 was proved in Ex. 3, and formula 10 can be proved in a similar manner.

Further transforms may be obtained by the following important property of the Laplace transformation.

Theorem 2 (First shifting theorem). If $\mathcal{L}(f) = F(s)$ when $s > \alpha$, then

$$\mathcal{L}\{e^{at}f(t)\} = F(s - a) \quad (s > \alpha + a);$$

that is, the substitution of $s - a$ for s in the transform corresponds to the multiplication of the original function by e^{at} .

Proof. By definition,

$$F(s) = \int_0^\infty e^{-st} f(t) dt$$

and, therefore,

$$F(s - a) = \int_0^\infty e^{-(s-a)t} f(t) dt = \int_0^\infty e^{-st} [e^{at} f(t)] dt = \mathcal{L}\{e^{at}f(t)\}.$$

Example 4. By applying Theorem 2 to the formulas 4, 7, and 8 in Table 16 we obtain the following results.

$f(t)$	$\mathcal{L}(f)$
$e^{at}t^n$	$\frac{n!}{(s - a)^{n+1}}$
$e^{at} \cos \omega t$	$\frac{s - a}{(s - a)^2 + \omega^2}$
$e^{at} \sin \omega t$	$\frac{\omega}{(s - a)^2 + \omega^2}$

We shall now state simple sufficient conditions for the existence of the Laplace transform of a function. These conditions are that f is piecewise continuous and $|f|$ does not increase too rapidly as t approaches infinity. We first state the definition of piecewise continuity.

A function $f(t)$ is said to be **piecewise continuous** on a finite interval $a \leq t \leq b$, if it is defined on that interval and is such that the interval can be subdivided into finitely many intervals, in each of which $f(t)$ is continuous and has finite limits as t approaches either end point of the interval of subdivision from the interior.

It follows from the definition that finite jumps are the only discontinuities which a piecewise continuous function may have; these are known as **ordinary discontinuities**. Furthermore, it is clear that the class of piecewise continuous functions includes every continuous function.

Theorem 3 (Existence theorem). Let $f(t)$ be a function which is piecewise continuous on every finite interval in the range $t \geq 0$ and satisfies

$$(2) \quad |f(t)| \leq Me^{\alpha t} \quad \text{for all } t \geq 0$$

and for some constants α and M . Then the Laplace transform of $f(t)$ exists for all $s > \alpha$.

Proof. Since $f(t)$ is piecewise continuous, $e^{-st}f(t)$ is integrable over any finite interval on the t -axis, and from (2),

$$\begin{aligned} |\mathcal{L}(f)| &= \left| \int_0^\infty e^{-st}f(t) dt \right| \leq \int_0^\infty e^{-st} |f(t)| dt \leq \int_0^\infty e^{-st} M e^{\alpha t} dt \\ &= M \int_0^\infty e^{-(s-\alpha)t} dt = \frac{M}{s-\alpha} \quad (s > \alpha). \end{aligned}$$

This completes the proof.

The conditions in Theorem 3 are practical for most applications, and it is simple to find whether a given function satisfies an inequality of the form (2). For example,

$$(3) \quad \cosh t < e^t, \quad t^n < n!e^t \quad (n = 0, 1, \dots) \quad \text{for all } t > 0,$$

and any function which is bounded in absolute value for all $t \geq 0$, such as the sine and cosine function of a real variable, satisfies that condition. An example of a function which does not satisfy a relation of the form (2), is the exponential function e^t , because, no matter how large we choose the numbers M and α in (2),

$$e^t > M e^{\alpha t} \quad \text{for all } t > t_0$$

where t_0 is a sufficiently large number, depending on M and α .

It should be noted that the conditions in Theorem 3 are sufficient rather than necessary. For example, the function $1/\sqrt{t}$ is infinite at $t = 0$, but its transform exists; in fact, from the definition and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ [cf. (10) in Sec. 17.1] we obtain

$$\mathcal{L}(t^{-\frac{1}{2}}) = \int_0^\infty e^{-st} t^{-\frac{1}{2}} dt = \frac{1}{\sqrt{s}} \int_0^\infty e^{-x} x^{-\frac{1}{2}} dx = \frac{1}{\sqrt{s}} \Gamma(\frac{1}{2}) = \sqrt{\frac{\pi}{s}}.$$

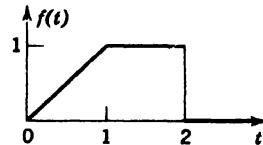
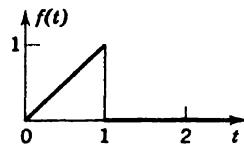
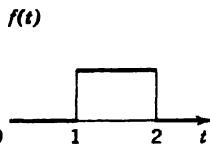
If the Laplace transform of a given function exists, it is uniquely determined. Conversely, it can be shown that two functions having the same transform cannot differ over an interval of positive length, although they may differ at various isolated points (cf. Ref. [B10] in Appendix 1). Since this is of no importance in applications, we may say that the inverse of a given transform is essentially unique. In particular, if two *continuous* functions have the same transform, they are completely identical.

Problems

- Extend Theorem 1 to linear combinations of more than two functions.
- Prove (3).

Find the Laplace transforms of the following functions, where a , b , etc., are constants.

- | | | | |
|-----------------------|-----------------------------|---------------------------------|--|
| 3. $t^2 - 2$ | 4. $a + bt + ct^2$ | 5. $4t^3 + t^2$ | 6. e^{at+b} |
| 7. $a \cos 2t$ | 8. $2 \cosh 4t$ | 9. $\cos^2 t$ | 10. $\sin^2 2t$ |
| 11. $e^{-2t} \cos t$ | 12. $e^{-at} \sin \omega t$ | 13. $2 \cos(\omega t + \theta)$ | 14. $Ce^{-at} \sin(\omega t + \theta)$ |
| 15. $e^{-t} \cos^2 t$ | 16. $e^{-t} \sin^2 t$ | 17. $3 \cosh^2 2t$ | 18. $\sinh^2 at$ |
| 19. | 20. | 21. | |



Show that

$$22. \mathcal{L}(te^{at}) = \frac{1}{(s-a)^2}$$

$$23. \mathcal{L}(t^n e^{at}) = \frac{n!}{(s-a)^{n+1}}$$

$$24. \mathcal{L}(e^{at} \cos \omega t) = \frac{s-a}{(s-a)^2 + \omega^2}$$

$$25. \mathcal{L}(e^{at} \sin \omega t) = \frac{\omega}{(s-a)^2 + \omega^2}$$

Representing the hyperbolic functions in terms of exponential functions and applying Theorem 2, show that

$$26. \mathcal{L}(\cosh at \cos at) = \frac{s^3}{s^4 + 4a^4}$$

$$27. \mathcal{L}(\cosh at \sin at) = \frac{a(s^2 + 2a^2)}{s^4 + 4a^4}$$

$$28. \mathcal{L}(\sinh at \cos at) = \frac{a(s^2 - 2a^2)}{s^4 + 4a^4}$$

$$29. \mathcal{L}(\sinh at \sin at) = \frac{2a^2 s}{s^4 + 4a^4}$$

Find $f(t)$ if $\mathcal{L}(f)$ equals

$$30. \frac{3}{s+2}$$

$$31. \frac{1}{s^2+9}$$

$$32. \frac{2s+1}{s^2-4}$$

$$33. \frac{3s-2}{s^2+1}$$

~~34. $\frac{1}{s^n}$ ($n = 1, 2, \dots$)~~

$$35. \frac{s+2}{(s+2)^2 + \omega^2} \quad 36. \frac{1}{(s-3)^2}$$

$$37. \frac{1}{(s-a)(s-b)} \quad (a \neq b)$$

$$38. \frac{a_1}{s} + \frac{a_2}{s^2} + \dots + \frac{a_n}{s^n}$$

39. Obtain formula 6 in Table 16 from formula 1 and Theorem 2.

40. Prove formula 4 in Table 16 by induction.

41. Obtain formula 10 in Table 16 from formula 6.

42. Obtain formula 6 in Table 16 from the formulas 9 and 10.

43. Derive the formulas 7 and 8 in Table 16 by integration by parts.

44. Using $\cosh x = \cos(ix)$ and $\sinh x = -i \sin(ix)$, $i = \sqrt{-1}$, obtain the formulas 9 and 10 in Table 16 from the formulas 7 and 8.

45. Using $\cos x = \cosh(ix)$ and $\sin x = -i \sinh(ix)$, derive the formulas 7 and 8 in Table 16 from the formulas 9 and 10.

4.2 Laplace Transforms of Derivatives and Integrals

We shall now see that, roughly speaking, differentiation and integration of $f(t)$ corresponds to multiplication and division of the Laplace transform $F(s) = \mathcal{L}(f)$ by s . The great importance of this property of the Laplace transformation is obvious, because in this way operations of calculus may be replaced by simple algebraic operations on the transforms. We start with the following theorem.

Theorem 1 [Differentiation of $f(t)$]. Suppose that $f(t)$ is continuous for all $t \geq 0$, satisfies (2), p. 196, for some α and M , and has a derivative $f'(t)$ which is piecewise continuous on every finite interval in the range $t \geq 0$. Then the Laplace transform of the derivative $f'(t)$ exists when $s > \alpha$, and

$$(1) \quad \mathcal{L}(f') = s\mathcal{L}(f) - f(0) \quad (s > \alpha).$$

Proof. We first consider the case when $f'(t)$ is continuous for all $t \geq 0$. Then, by definition and by integrating by parts,

$$\mathcal{L}(f') = \int_0^\infty e^{-st} f'(t) dt = [e^{-st} f(t)]_0^\infty + s \int_0^\infty e^{-st} f(t) dt.$$

Since f satisfies (2), Sec. 4.1, the integrated portion on the right is zero at the upper limit when $s > \alpha$, and at the lower limit it is $-f(0)$. The last integral is $\mathcal{L}(f)$, the existence for $s > \alpha$ being a consequence of Theorem 3 in Sec. 4.1. This proves that the expression on the right exists when $s > \alpha$, and is equal to $-f(0) + s\mathcal{L}(f)$. Consequently, $\mathcal{L}(f')$ exists when $s > \alpha$, and (1) holds.

If $f'(t)$ is merely piecewise continuous, the proof is quite similar; in this case, the range of integration in the original integral must be broken up into parts such that f' is continuous in each such part.

This theorem may be extended to piecewise continuous functions $f(t)$, but in place of (1) we then obtain the formula (1*) in Prob. 21 at the end of the section.

By applying (1) to the second-order derivative $f''(t)$ we obtain

$$\begin{aligned} \mathcal{L}(f'') &= s\mathcal{L}(f') - f'(0) \\ &= s[s\mathcal{L}(f) - f(0)] - f'(0); \end{aligned}$$

that is,

$$(2) \quad \mathcal{L}(f'') = s^2\mathcal{L}(f) - sf(0) - f'(0).$$

Similarly,

$$(3) \quad \mathcal{L}(f''') = s^3\mathcal{L}(f) - s^2f(0) - sf'(0) - f''(0),$$

etc. By induction we thus obtain the following extension of Theorem 1.

Theorem 2 (Derivative of any order n). Let $f(t)$ and its derivatives $f'(t)$, $f''(t)$, \dots , $f^{(n-1)}(t)$ be continuous functions for all $t \geq 0$, satisfying (2), p. 196, for some α and M , and let the derivative $f^{(n)}(t)$ be piecewise continuous on every finite interval in the range $t \geq 0$. Then the Laplace transform of $f^{(n)}(t)$ exists when $s > \alpha$, and is given by the formula

$$(4) \quad \mathcal{L}(f^{(n)}) = s^n \mathcal{L}(f) - s^{n-1}f(0) - s^{n-2}f'(0) - \dots - f^{(n-1)}(0).$$

An obvious application of Theorems 1 and 2 will be the transformation of derivatives in differential equations; this problem will be considered in Sec. 4.3. However, these theorems may also be used for determining transforms; cf. the following examples.

Example 1. Let $f(t) = t^2/2$. Find $\mathcal{L}(f)$. We have $f(0) = 0$, $f'(0) = 0$, $f''(t) = 1$.

Since $\mathcal{L}(1) = 1/s$, we obtain from (2)

$$\mathcal{L}(f'') = \mathcal{L}(1) = \frac{1}{s} = s^2 \mathcal{L}(f) \quad \text{or} \quad \mathcal{L}\left(\frac{t^2}{2}\right) = \frac{1}{s^3}.$$

Example 2. Let $f(t) = \cos \omega t$. Find $\mathcal{L}(f)$. We have $f(0) = 1$, and furthermore

$$f'(t) = -\omega \sin \omega t, \quad f'(0) = 0, \quad f''(t) = -\omega^2 \cos \omega t = -\omega^2 f(t).$$

From this and (2) it follows that

$$-\omega^2 \mathcal{L}(f) = \mathcal{L}(f'') = s^2 \mathcal{L}(f) - s \quad \text{or} \quad \mathcal{L}(f)(s^2 + \omega^2) = s.$$

Hence we obtain the expected result

$$\mathcal{L}(\cos \omega t) = \frac{s}{s^2 + \omega^2}.$$

Example 3. Let $f(t) = t \sin \omega t$. Find $\mathcal{L}(f)$. We have $f(0) = 0$,

$$f'(t) = \sin \omega t + \omega t \cos \omega t, \quad f'(0) = 0,$$

and furthermore

$$f''(t) = 2\omega \cos \omega t - \omega^2 t \sin \omega t = 2\omega \cos \omega t - \omega^2 f(t).$$

From this and (2) it follows that

$$2\omega \mathcal{L}(\cos \omega t) - \omega^2 \mathcal{L}(f) = \mathcal{L}(f'') = s^2 \mathcal{L}(f).$$

Using the formula for the Laplace transform of $\cos \omega t$, we thus obtain

$$(s^2 + \omega^2) \mathcal{L}(f) = 2\omega \mathcal{L}(\cos \omega t) = \frac{2\omega s}{s^2 + \omega^2}.$$

Consequently, the result is

$$\mathcal{L}(t \sin \omega t) = \frac{2\omega s}{(s^2 + \omega^2)^2}.$$

We have seen that differentiation of $f(t)$ corresponds to multiplication of $\mathcal{L}(f)$ by s (and addition of a constant). We shall now prove that integration of $f(t)$ corresponds to division of $\mathcal{L}(f)$ by s .

Theorem 3 [Integration of $f(t)$]. If $f(t)$ is piecewise continuous and satisfies an inequality of the form (2), Sec. 4.1, then

$$(5) \quad \mathcal{L}\left\{\int_0^t f(\tau) d\tau\right\} = \frac{1}{s} \mathcal{L}\{f(t)\} \quad (s > 0, s > \alpha).$$

Proof. Suppose that $f(t)$ is piecewise continuous and satisfies (2), Sec. 4.1, for some α and M . Clearly, if (2) holds for some negative α , it also holds for positive α , and we may assume that α is positive. Then the integral

$$g(t) = \int_0^t f(\tau) d\tau$$

is continuous, and by using (2) we obtain

$$|g(t)| \leq \int_0^t |f(\tau)| d\tau \leq M \int_0^t e^{\alpha\tau} d\tau = \frac{M}{\alpha} (e^{\alpha t} - 1) \quad (\alpha > 0).$$

Furthermore, $g'(t) = f(t)$, except for points at which $f(t)$ is discontinuous. Hence $g'(t)$ is piecewise continuous on each finite interval, and, according to Theorem 1,

$$\mathcal{L}\{f(t)\} = \mathcal{L}\{g'(t)\} = s\mathcal{L}\{g(t)\} - g(0) \quad (s > \alpha).$$

Clearly, $g(0) = 0$, and, therefore,

$$\mathcal{L}(g) = \frac{1}{s} \mathcal{L}(f).$$

This completes the proof.

Example 4. Let $\mathcal{L}(f) = \frac{1}{s^2(s^2 + \omega^2)}$. Find $f(t)$. From Table 16 in Sec. 4.1 we have

$$\mathcal{L}^{-1}\left(\frac{1}{s^2 + \omega^2}\right) = \frac{1}{\omega} \sin \omega t.$$

From this and Theorem 3 it follows that

$$\mathcal{L}^{-1}\left(\frac{1}{s}\left(\frac{1}{s^2 + \omega^2}\right)\right) = \frac{1}{\omega} \int_0^t \sin \omega \tau d\tau = \frac{1}{\omega^2} (1 - \cos \omega t).$$

Applying Theorem 3 once more, we obtain the desired answer

$$\mathcal{L}^{-1}\left(\frac{1}{s^2}\left(\frac{1}{s^2 + \omega^2}\right)\right) = \frac{1}{\omega^2} \int_0^t (1 - \cos \omega \tau) d\tau = \frac{1}{\omega^2} \left(t - \frac{\sin \omega t}{\omega}\right).$$

Problems

Using (1), derive

1. $\mathcal{L}(\sin t)$ from $\mathcal{L}(\cos t)$

2. $\mathcal{L}(\sinh 2t)$ from $\mathcal{L}(\cosh 2t)$

Using Theorems 1 and 2, show that

3. $\mathcal{L}(t \cos \omega t) = \frac{s^2 - \omega^2}{(s^2 + \omega^2)^2}$

4. $\mathcal{L}(t \cosh \omega t) = \frac{s^2 + \omega^2}{(s^2 - \omega^2)^2}$

5. $\mathcal{L}(t \sinh \omega t) = \frac{2as}{(s^2 - \omega^2)^2}$

6. $\mathcal{L}(te^{\alpha t}) = \frac{1}{(s - a)^2}$

Using Ex. 3 and Prob. 3, show that

$$\cdot 7. \mathcal{L}^{-1}\left(\frac{1}{(s^2 + \omega^2)^2}\right) = \frac{1}{2\omega^3} (\sin \omega t - \omega t \cos \omega t)$$

$$8. \mathcal{L}^{-1}\left(\frac{s^2}{(s^2 + \omega^2)^2}\right) = \frac{1}{2\omega} (\sin \omega t + \omega t \cos \omega t)$$

9. Find $\mathcal{L}(f)$ in Prob. 4 from Prob. 3 and $\cosh x = \cos(ix)$.

$$10. \text{Find } f(t) \text{ in Ex. 4 by noting that } \mathcal{L}(f) = \frac{1}{\omega^2} \left(\frac{1}{s^2} - \frac{1}{s^2 + \omega^2} \right).$$

Using Theorem 3, find $f(t)$ if $\mathcal{L}(f)$ equals

$$\checkmark 11. \frac{1}{s(s-2)}$$

$$\checkmark 12. \frac{1}{s(s^2+9)}$$

$$13. \frac{1}{s(s^2-1)}$$

$$14. \frac{1}{s^2(s+1)}$$

$$15. \frac{1}{s^2} \left(\frac{s-1}{s+1} \right)$$

$$16. \frac{1}{s^2} \left(\frac{s-2}{s^2+4} \right)$$

$$17. \frac{54}{s^3(s-3)}$$

$$18. \frac{2s-\pi}{s^3(s-\pi)}$$

$$19. \frac{6\pi^5}{s^4(s^2+\pi^2)}$$

20. Carry out the details of the proof of Theorem 1, assuming that $f'(t)$ has finite jumps at t_1, t_2, \dots, t_m , where $0 < t_1 < t_2 < \dots < t_m$.

21. Show that if $f(t)$ is continuous, except for an ordinary discontinuity at $t = a (> 0)$, the other conditions remaining the same as in Theorem 1, then

$$(1^*) \quad \mathcal{L}(f') = s\mathcal{L}(f) - f(0) - [f(a+0) - f(a-0)]e^{-as}.$$

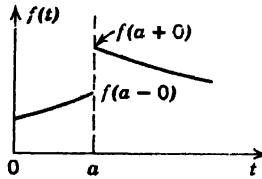


Fig. 82. Problem 21.

Graph the following functions. Using (1*), find their Laplace transform.

$$22. f(t) = 1 \text{ when } 1 < t < 2, \quad f(t) = 0 \text{ otherwise}$$

$$23. f(t) = t \text{ when } 0 < t < 1, \quad f(t) = 0 \text{ otherwise}$$

$$24. f(t) = t \text{ when } 0 < t < 1, \quad f(t) = 1 \text{ when } 1 < t < 2, \quad f(t) = 0 \text{ otherwise}$$

$$25. f(t) = t - 1 \text{ when } 1 < t < 2, \quad f(t) = 0 \text{ otherwise}$$

4.3 Transformation of Ordinary Differential Equations

We shall now see that ordinary linear differential equations with constant coefficients can be reduced to algebraic equations of the transform. For example, consider the equation

$$(1) \quad y''(t) + \omega^2 y(t) = r(t)$$

where $r(t)$ and ω are given. Applying the Laplace transformation and using (2), Sec. 4.2, we obtain

$$(2) \quad s^2 Y(s) - sy(0) - y'(0) + \omega^2 Y(s) = R(s),$$

where $Y(s)$ is the Laplace transform of the (unknown) function $y(t)$, and $R(s)$ is the Laplace transform of $r(t)$. This algebraic equation is called the **subsidiary equation** of the given differential equation. Its solution is clearly

$$(3) \quad Y(s) = \frac{sy(0) + y'(0)}{s^2 + \omega^2} + \frac{R(s)}{s^2 + \omega^2}.$$

Note that the first term on the right is completely determined by means of given initial conditions, $y(0) = k_1$, $y'(0) = k_2$.

The last step of the procedure is to determine the inverse $\mathcal{L}^{-1}(Y) = y(t)$ which is then the desired solution of (1).

Note that (2) was obtained from (1) by means of Theorem 2, Sec. 4.2. Instead of justifying this step by showing that the final result $y(t)$ satisfies the conditions of that theorem, it is simpler and more natural to see by substitution whether $y(t)$ satisfies the given equation and the initial conditions.

Systems of differential equations will be considered in Sec. 4.5.

Example 1. Find the solution of the differential equation

$$y'' + 9y = 0,$$

satisfying the initial conditions $y(0) = 0$, $y'(0) = 2$.

From (2) it follows that the subsidiary equation is

$$s^2 Y(s) - 2 + 9Y(s) = 0.$$

Solving for $Y(s)$, we obtain

$$Y(s) = \frac{2}{s^2 + 9} = \frac{2}{3} \left(\frac{3}{s^2 + 9} \right).$$

From this and Table 16 in Sec. 4.1 we find

$$y(t) = \mathcal{L}^{-1}(Y) = \frac{2}{3} \sin 3t.$$

By substitution we see that this function satisfies the given equation and the initial conditions.

Example 2. A small body of mass $m = 2$ is attached at the lower end of an elastic spring whose upper end is fixed, the spring modulus being $k = 10$. Let $y(t)$ be the displacement of the body from the position of static equilibrium. Determine the free vibrations of the body, starting from the initial position $y(0) = 2$ with the initial velocity $y'(0) = -4$, assuming that there is damping proportional to the velocity, the damping constant being $c = 4$.

The motion is described by the solution $y(t)$ of the initial value problem

$$y'' + 2y' + 5y = 0, \quad y(0) = 2, \quad y'(0) = -4,$$

cf. (7), Sec. 2.6. Using (1) and (2) in Sec. 4.2, we obtain the subsidiary equation

$$s^2 Y - 2s + 4 + 2(sY - 2) + 5Y = 0.$$

The solution is

$$Y(s) = \frac{2s}{(s+1)^2 + 2^2} = 2 \frac{s+1}{(s+1)^2 + 2^2} - \frac{2}{(s+1)^2 + 2^2}.$$

Now

$$\mathcal{L}^{-1}\left(\frac{s}{s^2 + 2^2}\right) = \cos 2t, \quad \mathcal{L}^{-1}\left(\frac{2}{s^2 + 2^2}\right) = \sin 2t.$$

From this and Theorem 2 in Sec. 4.1 we obtain the expected type of solution

$$y(t) = \mathcal{L}^{-1}(Y) = e^{-t}(2 \cos 2t - \sin 2t).$$

Problems

1. Using the Laplace transformation, show that the initial value problem

$$y'' + \omega^2 y = 0, \quad y(0) = A, \quad y'(0) = B \quad (\omega \text{ real, not zero})$$

has the solution $y(t) = A \cos \omega t + (B/\omega) \sin \omega t$.

2. Using the Laplace transformation, show that the initial value problem

$$y'' - a^2 y = 0, \quad y(0) = k_1, \quad y'(0) = k_2 \quad (a \text{ real, not zero})$$

has the solution $y(t) = k_1 \cosh at + \frac{k_2}{a} \sinh at$.

Using the Laplace transformation, solve the following initial value problems.

3. $y'' + 4y = 0, \quad y(0) = 1, \quad y'(0) = 1$
 4. $4y'' + \pi^2 y = 0, \quad y(0) = 0, \quad y'(0) = -\frac{1}{2}$
 5. $y'' - 16y = 0, \quad y(0) = -2, \quad y'(0) = 0$
 6. $25y'' - 9y = 0, \quad y(0) = 1, \quad y'(0) = 0$

7. Show that the subsidiary equation of the differential equation

$$y'' + ay' + by = r(t) \quad (a, b \text{ constant})$$

has the solution

$$Y(s) = \frac{(s + a)y(0) + y'(0)}{s^2 + as + b} + \frac{R(s)}{s^2 + as + b}$$

where $R(s)$ is the Laplace transform of $r(t)$.

Using the Laplace transformation, solve the following initial value problems.

8. $y'' + y' - 2y = 0, \quad y(0) = 0, \quad y'(0) = 3$
 9. $y'' - 2y' - 3y = 0, \quad y(0) = 1, \quad y'(0) = 7$
 10. $y'' + 2y' - 8y = 0, \quad y(0) = 1, \quad y'(0) = 8$
 11. $y'' + 2y' - 3y = 0, \quad y(0) = 0, \quad y'(0) = 4$
 12. $4y'' + 4y' - 3y = 0, \quad y(0) = 8, \quad y'(0) = 0$
 13. $y'' - 4y' + 5y = 0, \quad y(0) = 1, \quad y'(0) = 3$
 14. $y'' + 2y' + 5y = 0, \quad y(0) = 2, \quad y'(0) = -4$
 15. $y'' + 2y' + y = 0, \quad y(0) = 1, \quad y'(0) = -2$
 16. $y'' - 4y' + 4y = 0, \quad y(0) = 0, \quad y'(0) = 2$
 17. $4y'' - 4y' + 37y = 0, \quad y(0) = 1, \quad y'(0) = 0.5$
 18. $y'' + 2y' + 10y = 0, \quad y(0) = 0, \quad y'(0) = 3$
 19. $9y'' - 6y' + y = 0, \quad y(0) = 3, \quad y'(0) = 1$
 20. $y'' + 4y' + 4y = 0, \quad y(0) = 2, \quad y'(0) = -3$

4.4 Partial Fractions

Our previous consideration shows that in many cases of practical importance the solution $Y(s)$ of the subsidiary equation of a given differential equation will be of the form

$$Y(s) = \frac{G(s)}{H(s)}$$

where $G(s)$ and $H(s)$ are polynomials of s . We shall now see that in such a case we can determine $y(t) = \mathcal{L}^{-1}(Y)$ by expressing $Y(s)$ in terms of partial fractions. Throughout this section we shall make the following

Assumption. $G(s)$ and $H(s)$ have no common factors and have real coefficients. The degree of $G(s)$ is lower than that of $H(s)$.

Let $s = a$ be a root of $H(s) = 0$. We shall state formulas for the inverse transform of the partial fraction or fractions corresponding to a in the representation of $Y(s)$ in terms of partial fractions. The proofs will be given below and examples in the next section.

Notation. (1a), (2a), ..., (5a) (see below) are representations of $Y(s) = G(s)/H(s)$, and in each case $W(s)$ denotes the sum of the partial fractions corresponding to all the (unrepeated or repeated) linear factors of $H(s)$ which are not under consideration.

Case 1. Unrepeated factor $s - a$. In this case,

$$(1a) \quad Y(s) = \frac{G(s)}{H(s)} = \frac{A}{s-a} + W(s).$$

The inverse transform is

$$(1b) \quad \mathcal{L}^{-1}(Y) = Ae^{at} + \mathcal{L}^{-1}(W)$$

where A is given by either of the two expressions

$$(1c) \quad A = Q_a(a) \quad \text{or} \quad A = \frac{G(a)}{H'(a)}$$

and $Q_a(s)$ is the function which is left after removing the factor $s - a$ from $H(s)$ in $Y(s)$, that is

$$(1d) \quad Q_a(s) = \frac{(s-a)G(s)}{H(s)}.$$

The subscript a of Q_a emphasizes the important fact that Q_a depends on a and changes when we proceed from one linear factor to another.

Case 2. Repeated factor $(s - a)^m$. In this case,

$$(2a) \quad Y(s) = \frac{G(s)}{H(s)} = \frac{A_m}{(s - a)^m} + \frac{A_{m-1}}{(s - a)^{m-1}} + \cdots + \frac{A_1}{s - a} + W(s).$$

The inverse transform of this expression is

$$(2b) \quad \mathcal{L}^{-1}(Y) = e^{at} \left(A_m \frac{t^{m-1}}{(m-1)!} + A_{m-1} \frac{t^{m-2}}{(m-2)!} + \cdots + A_2 \frac{t}{1!} + A_1 \right) + \mathcal{L}^{-1}(W)$$

where the constants A_1, \dots, A_m are given by the formula

$$(2c) \quad A_m = Q_a(a), \quad A_k = \frac{1}{(m-k)!} \left. \frac{d^{m-k} Q_a(s)}{ds} \right| \quad (k = 1, 2, \dots, m-1);$$

here,

$$(2d) \quad Q_a(s) = \frac{(s-a)^m G(s)}{H(s)}.$$

Complex factors. The previous cases include real as well as complex numbers a , but if a is complex, other formulas are preferable from the practical point of view. Let

$$a = \alpha + i\beta \quad \text{and} \quad \bar{a} = \alpha - i\beta \quad (\alpha, \beta \text{ real}, \beta \neq 0).$$

Suppose that $s = a$ is a root of $H(s) = 0$. Then since $H(s)$ has real coefficients, $s = \bar{a}$ is also a root of $H(s) = 0$.

Case 3. Unrepeated complex factor $s - a$. Writing the partial fraction corresponding to $s - a$ and $s - \bar{a}$ explicitly and noting that

$$(s - a)(s - \bar{a}) = (s - \alpha)^2 + \beta^2,$$

we have

$$(3a) \quad Y(s) = \frac{G(s)}{H(s)} = \frac{As + B}{(s - \alpha)^2 + \beta^2} + W(s) \quad (A, B \text{ real}).$$

The inverse transform is

$$(3b) \quad \mathcal{L}^{-1}(Y) = \frac{1}{\beta} e^{at} (T_a \cos \beta t + S_a \sin \beta t) + \mathcal{L}^{-1}(W)$$

where $\overline{S_a}$ and T_a are the real and imaginary parts of

$$(3c) \quad R_a(a) = S_a + iT_a;$$

here

$$(3d) \quad R_a(s) = [(s - \alpha)^2 + \beta^2] \frac{G(s)}{H(s)}.$$

Case 4. Repeated complex factor $(s - a)^2$. Writing the partial fractions corresponding to $(s - a)^2$ and $(s - \bar{a})^2$ explicitly, we have

$$(4a) \quad Y(s) = \frac{G(s)}{H(s)} = \frac{As + B}{[(s - \alpha)^2 + \beta^2]^2} + \frac{Cs + D}{(s - \alpha)^2 + \beta^2} + W(s)$$

where A, B, C, D are real. The inverse transform is

$$(4b) \quad \mathcal{L}^{-1}(Y) = \frac{1}{2\beta^3} e^{at} [(T_a - \beta S_a^* - \beta S_a t) \cos \beta t + (S_a + \beta T_a^* + \beta T_a t) \sin \beta t] + \mathcal{L}^{-1}(W)$$

where

$$(4c) \quad R_a(a) = S_a + iT_a, \quad R_a'(a) = S_a^* + iT_a^*$$

and

$$(4d) \quad R_a(s) = [(s - \alpha)^2 + \beta^2]^2 \frac{G(s)}{H(s)}.$$

Case 5. Repeated complex factor $(s - a)^m$. Writing the partial fractions corresponding to $(s - a)^m$ and $(s - \bar{a})^m$ explicitly, we have

$$(5a) \quad Y(s) = \frac{G(s)}{H(s)} = \sum_{k=1}^m \left\{ \frac{A_k}{(s - a)^k} + \frac{B_k}{(s - \bar{a})^k} \right\} + W(s).$$

The transform is

$$(5b) \quad \mathcal{L}^{-1}(Y) = 2e^{at} \sum_{k=1}^m \frac{t^{k-1}}{(k-1)! (m-k)!} (Q_{k1} \cos \beta t - Q_{k2} \sin \beta t) + \mathcal{L}^{-1}(W)$$

where

$$(5c) \quad Q_a(a)^{(m-k)} = Q_{k1} + iQ_{k2} = \frac{d^{m-k} Q_a(s)}{ds^{m-k}} \Big|_{s=a}, \quad Q_a(a)^{(0)} = Q_a(a),$$

and

$$(5d) \quad Q_a(s) = \frac{(s - a)^m G(s)}{H(s)}.$$

Remark. When $m = 2$, the previous formulas are preferable. Case 5 with $m > 2$ is of minor practical importance, and we present those formulas merely for the sake of completeness.

Proofs of the formulas

Proof in Case 1. Multiplying (1a) by $s - a$ and using (1d), we have

$$Q_a(s) = A + (s - a)W(s).$$

Letting $s \rightarrow a$ and noting that $W(s)$ does not contain the factor $s - a$, we obtain $A = Q_a(a)$, the first formula in (1c). If we write (1d) in the form

$$Q_a(s) = \frac{G(s)}{H(s)/(s - a)}$$

and let $s \rightarrow a$, then $G(s) \rightarrow G(a)$ while the denominator appears as an indeterminate of the form $0/0$; evaluating it in the usual way, we obtain

$$\lim_{s \rightarrow a} \frac{H(s)}{s - a} = \lim_{s \rightarrow a} \frac{H'(s)}{(s - a)'} = H'(a).$$

This proves the second formula in (1c). The inverse transform of $A/(s - a)$ is Ae^{at} , and the proof is complete.

Proof in Case 2. Multiplying (2a) by $(s - a)^m$ and using (2d), we have

$$Q_a(s) = A_m + (s - a)A_{m-1} + (s - a)^2 A_{m-2} + \cdots + (s - a)^m W(s).$$

Hence $Q_a(a) = A_m$. Differentiation yields

$$Q_a'(s) = A_{m-1} + \text{further terms containing factors } s - a.$$

Hence $Q_a'(a) = A_{m-1}$ because those terms are zero when $s = a$. Differentiating again we find $Q_a''(a) = 2! A_{m-2}$, etc. This proves (2c). The inverse transform of $A_k/(s - a)^k$ is $A_k e^{at} t^{k-1}/(k - 1)!$. This yields (2b), and the proof is complete.

Proof in Case 3. Multiplying (3a) by $(s - \alpha)^2 + \beta^2$, using (3d), and letting $s \rightarrow a$, we obtain

$$R_a(a) = aA + B = (\alpha + i\beta)A + B.$$

The real and imaginary parts on both sides must be equal. Using (3c), we thus have

$$S_a = \alpha A + B, \quad T_a = \beta A.$$

Hence in (3a),

$$\frac{As + B}{(s - \alpha)^2 + \beta^2} = \frac{A(s - \alpha) + \alpha A + B}{(s - \alpha)^2 + \beta^2} = \frac{1}{\beta} \frac{T_a(s - \alpha) + \beta S_a}{(s - \alpha)^2 + \beta^2}.$$

From this, Table 16, and Theorem 2 in Sec. 4.1 we obtain (3b). This completes the proof.

Proof in Case 4. Introducing the notation

$$p(s) = (s - a)(s - \bar{a}) = (s - \alpha)^2 + \beta^2,$$

multiplying (4a) by p^2 , and using (4d), we have

$$R_a(s) = As + B + (Cs + D)p(s) + p^2(s)W(s).$$

Letting $s \rightarrow a$ and noting $p(a) = 0$, we obtain

$$(6) \quad R_a(a) = aA + B = (\alpha + i\beta)A + B.$$

By differentiation with respect to s we find

$$R_a'(s) = A + (Cs + D)p'(s) + \text{terms containing a factor } p(s).$$

Since $p(a) = 0$ and $p'(a) = 2(a - \alpha) = 2i\beta$, it follows that

$$(7) \quad R_a'(a) = A + 2i\beta[C(\alpha + i\beta) + D].$$

In (6) and (7) the real and imaginary parts on both sides must be equal. Using (4c) we thus obtain

$$(8) \quad S_a = \alpha A + B, \quad T_a = \beta A.$$

and furthermore

$$S_a^* = A - 2\beta^2 C, \quad T_a^* = 2\alpha\beta C + 2\beta D,$$

which yields

$$(9) \quad C = (T_a - \beta S_a^*)/2\beta^3, \quad \alpha C + D = T_a^*/2\beta.$$

From Table 16 and Theorem 2 in Sec. 4.1 we see that the partial fraction

$$\frac{Cs + D}{(s - \alpha)^2 + \beta^2} = \frac{C(s - \alpha) + \alpha C + D}{(s - \alpha)^2 + \beta^2}$$

in (4a) has the inverse transform

$$e^{\alpha t}[\beta C \cos \beta t + (D + \alpha C) \sin \beta t]/\beta.$$

The other partial fraction

$$\frac{As + B}{[(s - \alpha)^2 + \beta^2]^2} = \frac{A(s - \alpha) + \alpha A + B}{[(s - \alpha)^2 + \beta^2]^2}$$

has the inverse transform

$$e^{\alpha t}[(A\beta^2 t + \alpha A + B) \sin \beta t - (\alpha A + B)\beta t \cos \beta t]/2\beta^3,$$

as follows from Example 3 and Problem 7 in Sec. 4.2 and Theorem 2 in Sec. 4.1. By inserting (8) and (9) and simplifying we obtain (4b), and the proof is complete.

Proof in Case 5. From (2b) we see that the inverse transform of $A_k(s - a)^{-k} + B_k(s - \bar{a})^{-k}$ in (5a) is

$$(10) \quad \begin{aligned} & \frac{t^{k-1}}{(k-1)!} (A_k e^{\alpha t} + B_k e^{\bar{\alpha} t}) \\ &= \frac{t^{k-1}}{(k-1)!} e^{\alpha t} [(A_k + B_k) \cos \beta t + i(A_k - B_k) \sin \beta t]. \end{aligned}$$

Using the notation in (5c), we have from (2c)

$$A_k = \frac{1}{(m-k)!} Q_a(a)^{(m-k)} = \frac{Q_{k1} + iQ_{k2}}{(m-k)!}.$$

Similarly, since $G(s)$ and $H(s)$ have real coefficients,

$$B_k = \frac{1}{(m-k)!} Q_{\bar{a}}(\bar{a})^{(m-k)} = \frac{Q_{k1} - iQ_{k2}}{(m-k)!}.$$

By addition and subtraction we find

$$A_k + B_k = 2Q_{k1}/(m-k)!, \quad A_k - B_k = 2iQ_{k2}/(m-k)!.$$

If we insert this into (10), we obtain (5b), and the proof is complete.

4.5 Examples and Applications

Example 1. Find the inverse transformation of

$$Y(s) = \frac{G(s)}{H(s)} = \frac{s+1}{s^3 + s^2 - 6s}.$$

The denominator has the three distinct linear factors s , $s - 2$, and $s + 3$, corresponding to the roots $a_1 = 0$, $a_2 = 2$, and $a_3 = -3$. Hence

$$Y(s) = \frac{s+1}{s(s-2)(s+3)} = \frac{A_1}{s} + \frac{A_2}{s-2} + \frac{A_3}{s+3},$$

and $G(s) = s + 1$, $H'(s) = 3s^2 + 2s - 6$. Using (1c) in Sec. 4.4 we thus obtain

$$A_1 = \frac{G(0)}{H'(0)} = -\frac{1}{6}, \quad A_2 = \frac{G(2)}{H'(2)} = \frac{3}{10}, \quad A_3 = \frac{G(-3)}{H'(-3)} = -\frac{2}{15}.$$

From this and (1b) we see that the inverse transform is

$$\mathcal{L}^{-1}(Y) = -\frac{1}{6} + \frac{3}{10}e^{2t} - \frac{2}{15}e^{-3t}.$$

Example 2. Solve the initial value problem

$$y'' - 3y' + 2y = 4t + e^{3t}, \quad y(0) = 1, \quad y'(0) = -1.$$

From Table 16, Sec. 4.1, and (1) and (2) in Sec. 4.2 we see that the subsidiary equation is

$$s^2Y - s + 1 - 3(sY - 1) + 2Y = \frac{4}{s^2} + \frac{1}{s-3}.$$

Solving for Y and representing Y in terms of partial fractions, we have

$$\begin{aligned} Y(s) &= \frac{G(s)}{H(s)} = \frac{s^4 - 7s^3 + 13s^2 + 4s - 12}{s^2(s-3)(s^2-3s+2)} \\ &= \frac{A_2}{s^2} + \frac{A_1}{s} + \frac{B}{s-3} + \frac{C}{s-2} + \frac{D}{s-1}. \end{aligned}$$

We first determine the constants A_1 and A_2 in the partial fractions corresponding to the root $a = 0$. In (2d) and (2c), Sec. 4.4,

$$Q_0(s) = \frac{G(s)}{(s-3)(s^2-3s+2)}, \quad A_2 = Q_0(0) = 2, \quad A_1 = Q_0'(0) = 3.$$

Using (1c) we obtain the remaining coefficients

$$\begin{aligned} B &= \left. \frac{G(s)}{s^2(s^2-3s+2)} \right|_{s=3} = \frac{1}{2}, & C &= \left. \frac{G(s)}{s^2(s-3)(s-1)} \right|_{s=2} = -2, \\ D &= \left. \frac{G(s)}{s^2(s-3)(s-2)} \right|_{s=1} = -\frac{1}{2}. \end{aligned}$$

From (1b) and (2b) we see that the solution is

$$y(t) = \mathcal{L}^{-1}(Y) = 2t + 3 + \frac{1}{2}e^{3t} - 2e^{2t} - \frac{1}{2}e^t.$$

The student may solve the problem by the classical method and convince himself that the present approach is much simpler and much more rapid.

Example 3 (Free oscillations). Determine the free damped oscillation corresponding to the initial value problem

$$y'' + 2y' + 5y = 0, \quad y(0) = 2, \quad y'(0) = -4.$$

The solution of the subsidiary equation is (cf. Ex. 2 in Sec. 4.3)

$$Y(s) = \frac{2s}{s^2 + 2s + 5} = \frac{2s}{(s - a)(s - \bar{a})}$$

where $a = \alpha + i\beta = -1 + 2i$, so that $\alpha = -1$ and $\beta = 2$. Furthermore, in (3d) and (3c),

$$R_a(s) = 2s, \quad R_a(a) = S_a + iT_a = 2a = 2(\alpha + i\beta) = -2 + 4i,$$

that is, $S_a = -2$ and $T_a = 4$. From this and (3b),

$$y(t) = \mathcal{L}^{-1}(Y) = e^{-t}(2 \cos 2t - \sin 2t).$$

Example 4 (Forced oscillations, resonance). Solve the initial value problem

$$my'' + ky = F_0 \sin pt, \quad y(0) = 0, \quad y'(0) = 0.$$

From Sec. 2.13 we know that this equation governs the forced oscillations of a body of mass m attached at the lower end of an elastic spring whose upper end is fixed (Fig. 83). k is the spring modulus, and $F_0 \sin pt$ is the driving force or input. Setting $\omega_0 = \sqrt{k/m}$, we may write the equation in the form

$$y'' + \omega_0^2 y = K \sin pt \quad \left(K = \frac{F_0}{m} \right).$$

The subsidiary equation is

$$(1) \quad s^2 Y + \omega_0^2 Y = K \frac{p}{s^2 + p^2}.$$

No resonance. If $p^2 \neq \omega_0^2$, the solution of (1) and its representation in terms of partial fractions is

$$(2) \quad Y(s) = \frac{Kp}{(s^2 + \omega_0^2)(s^2 + p^2)} = \frac{As + B}{s^2 + \omega_0^2} + \frac{Ms + N}{s^2 + p^2}.$$

We first determine A and B in the fraction corresponding to $a_1 = \alpha_1 + i\beta_1 = i\omega_0$. In (3c), Sec. 4.4,

$$R_{a_1}(a_1) = S_{a_1} + iT_{a_1} = \frac{Kp}{s^2 + p^2} \Big|_{s=a_1} = \frac{Kp}{p^2 - \omega_0^2}$$

so that $T_{a_1} = 0$. From this and (3b) we see that the inverse transform of the first fraction on the right-hand side of (2) is

$$\frac{Kp}{\omega_0(p^2 - \omega_0^2)} \sin \omega_0 t.$$

We determine M and N in the last fraction, which corresponds to $a_2 = \alpha_2 + i\beta_2 = ip$. In (3c) we now have

$$R_{a_2}(a_2) = S_{a_2} + iT_{a_2} = \frac{Kp}{s^2 + \omega_0^2} \Big|_{s=a_2} = \frac{Kp}{\omega_0^2 - p^2}.$$

Hence $T_{a_2} = 0$. From (3b) we see that the inverse transform of that fraction is

$$\frac{K}{-\omega_0^2} \sin pt.$$

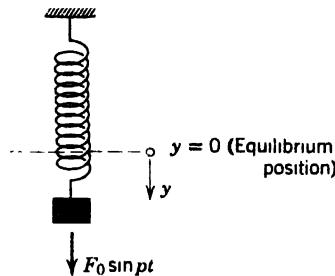


Fig. 83. Vibrating system. Forced oscillations.

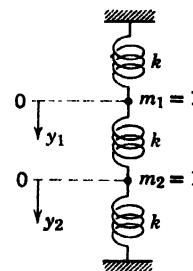


Fig. 84. Example 5.

By adding these two transforms we obtain the desired solution

$$y(t) = \mathcal{L}^{-1}(Y) = \frac{K}{p^2 - \omega_0^2} \left(\frac{p}{\omega_0} \sin \omega_0 t - \sin pt \right).$$

The vibration is a superposition of two harmonic motions, one of which has the frequency of the driving force and the other has the frequency of the freely vibrating system.

Resonance. If $p = \omega_0$, the solution of (1) is

$$Y(s) = \frac{K\omega_0}{(s^2 + \omega_0^2)^2}.$$

The denominator has the double roots $a = \alpha + i\beta = i\omega_0$ and $\bar{a} = -i\omega_0$, that is, $\alpha = 0$ and $\beta = \omega_0$. In (4c) and (4d), Sec. 4.4,

$$R_a(s) = K\omega_0, \quad S_a = K\omega_0, \quad T_a = 0, \quad R_a'(s) = 0,$$

and (4b) yields the solution

$$(3) \quad y(t) = \mathcal{L}^{-1}(Y) = \frac{K}{2\omega_0^2} (\sin \omega_0 t - \omega_0 t \cos \omega_0 t).$$

Systems of differential equations may also be solved by means of the Laplace transformation; cf. the following example.

Example 5 (Vibration of two masses). The vibrations of the mechanical system of two masses shown in Fig. 84 are governed by the differential equations

$$(4) \quad \begin{aligned} y_1'' &= -ky_1 + k(y_2 - y_1) \\ y_2'' &= -k(y_2 - y_1) - ky_2 \end{aligned}$$

where k is the spring modulus of each of the three springs, y_1 and y_2 are the displacements of the masses from their position of static equilibrium; the masses of the springs and the damping are neglected. The derivation of (4) is similar to that of the differential equation in Sec. 2.6.

We shall determine the solution corresponding to the initial conditions $y_1(0) = 1$, $y_2(0) = 1$, $y_1'(0) = \sqrt{3k}$, $y_2'(0) = -\sqrt{3k}$. Let $Y_1 = \mathcal{L}(y_1)$ and $Y_2 = \mathcal{L}(y_2)$. Then from (2) in Sec. 4.2 and the initial conditions we obtain the subsidiary equations

$$s^2 Y_1 - s - \sqrt{3k} = -k Y_1 + k(Y_2 - Y_1)$$

$$s^2 Y_2 - s + \sqrt{3k} = -k(Y_2 - Y_1) - k Y_2.$$

This system of linear algebraic equations in the unknowns Y_1 and Y_2 may be written

$$(s^2 + 2k)Y_1 - k Y_2 = s + \sqrt{3k}$$

$$-k Y_1 + (s^2 + 2k)Y_2 = s - \sqrt{3k}.$$

Cramer's rule (Sec. 0.3) yields the solution

$$Y_1 = \frac{(s + \sqrt{3k})(s^2 + 2k) + k(s - \sqrt{3k})}{(s^2 + 2k)^2 - k^2}$$

$$Y_2 = \frac{(s^2 + 2k)(s - \sqrt{3k}) + k(s + \sqrt{3k})}{(s^2 + 2k)^2 - k^2}.$$

The representations in terms of partial fractions are

$$(5) \quad Y_1 = \frac{s}{s^2 + k} + \frac{\sqrt{3k}}{s^2 + 3k}, \quad Y_2 = \frac{s}{s^2 + k} - \frac{\sqrt{3k}}{s^2 + 3k}.$$

Hence the solution of our initial value problem is

$$y_1(t) = \mathcal{L}^{-1}(Y_1) = \cos \sqrt{k}t + \sin \sqrt{3k}t$$

$$y_2(t) = \mathcal{L}^{-1}(Y_2) = \cos \sqrt{k}t - \sin \sqrt{3k}t.$$

Problems

Find $f(t)$, if $\mathcal{L}(f)$ equals

1. $\frac{3s - 2}{s^2 - s}$

2. $\frac{1}{(s - a)(s - b)}$

3. $\frac{s^2 + 9s - 9}{s^3 - 9s}$

4. $\frac{6}{(s^2 - 1)(s^2 - 4)}$

5. $\frac{11s - 14}{s^3 - s^2 - 4s + 4}$

6. $\frac{s - 2}{s^2 + 1}$

7. $\frac{4s + 4}{s^2 + 16}$

8. $\frac{s}{s^2 + 2s + 2}$

9. $\frac{4 - 2s}{s^3 + 4s}$

10. $\frac{3s^2 - 2s - 1}{(s - 3)(s^2 + 1)}$

11. $\frac{s}{(s + 1)^2}$

12. $\frac{s^2 - 4s}{(s - 2)^3}$

13. $\frac{s^3 + 6s^2 + 14s}{(s + 2)^4}$

14. $\frac{s^2 + 1}{s(s + 1)^2}$

15. $\frac{2s^2 - 3s}{(s - 2)(s - 1)^2}$

16. $\frac{3(s^3 - s + 3)}{(s - 1)^2(s + 2)^2}$

17. $\frac{s^2 + 2s}{(s^2 + 2s + 2)^2}$

18. $\frac{1 - s}{(s^2 - 2s + 2)^2}$

19. $\frac{s^2 - 6s + 7}{(s^2 - 4s + 5)^2}$

20. $\frac{3s^2 - 6s + 7}{(s^2 - 2s + 5)^2}$

21. $\frac{s^3 - 3s^2 + 6s - 4}{(s^2 - 2s + 2)^2}$

Using the Laplace transformation, solve the following initial value problems.

22. $y'' - y = 1, \quad y(0) = 1, \quad y'(0) = 2$

23. $y'' - 3y' + 2y = 6e^{-t}, \quad y(0) = 3, \quad y'(0) = 3$

24. $y'' + 2y' - 3y = 10 \sinh 2t, \quad y(0) = 0, \quad y'(0) = 4$

25. $y'' + 2y' + 2y = 0, \quad y(0) = 0, \quad y'(0) = 1$

26. $y'' + 4y' + 20y = 0, \quad y(0) = 1, \quad y'(0) = -10$

27. $y'' + 4y' + 5y = 5$, $y(0) = 1$, $y'(0) = 1$
 28. $y'' + 6y' + 13y = 0$, $y(0) = 2$, $y'(0) = -6$
 29. $y'' + 2y' + 2y = 10 \sin 2t$, $y(0) = -1$, $y'(0) = -3$
 30. $y'' - 2y' + 5y = 8 \sin t - 4 \cos t$, $y(0) = 1$, $y'(0) = 3$
 31. $y'' - 6y' + 9y = 4e^t$, $y(0) = 2$, $y'(0) = 4$
 32. $y'' + 2y' + y = 2 \cos t$, $y(0) = 3$, $y'(0) = 0$
 33. $y'' - 4y' + 4y = 8 \sin 2t$, $y(0) = 1$, $y'(0) = 1$
 34. $y'' + y = -2 \sin t$, $y(0) = 0$, $y'(0) = 1$
 35. $y'' + 9y = 6 \cos 3t$, $y(0) = 2$, $y'(0) = 0$
 36. $y'' + 4y = 4(\cos 2t - \sin 2t)$, $y(0) = 1$, $y'(0) = 3$

Show that

$$37. \mathcal{L}^{-1}\left\{\frac{s^3}{s^4 + 4a^4}\right\} = \cosh at \cos at \quad 38. \mathcal{L}^{-1}\left\{\frac{s}{s^4 + 4a^4}\right\} = \frac{1}{2a^2} \sinh at \sin at$$

$$39. \mathcal{L}^{-1}\left\{\frac{s^2}{s^4 + 4a^4}\right\} = \frac{1}{2a} (\cosh at \sin at + \sinh at \cos at)$$

$$40. \mathcal{L}^{-1}\left\{\frac{1}{s^4 + 4a^4}\right\} = \frac{1}{4a^3} (\cosh at \sin at - \sinh at \cos at)$$

41. Derive (4) in Ex. 5.

42. Applying the method of partial fractions, obtain (5).

Solve the following initial value problems involving systems of differential equations.

43. $y_1' = -y_2$, $y_2' = y_1$, $y_1(0) = 1$, $y_2(0) = 0$
 44. $y_1'' = y_1 + 3y_2$, $y_2'' = 4y_1 - 4e^t$,
 $y_1(0) = 2$, $y_1'(0) = 3$, $y_2(0) = 1$, $y_2'(0) = 2$
 45. $y_1'' + y_2 = -5 \cos 2t$, $y_2'' + y_1 = 5 \cos 2t$,
 $y_1(0) = 1$, $y_1'(0) = 1$, $y_2(0) = -1$, $y_2'(0) = 1$

4.6 Differentiation and Integration of Transforms

It can be shown that if $f(t)$ satisfies the conditions of the existence theorem in Sec. 4.1, then the derivative of the corresponding transform

$$F(s) = \mathcal{L}(f) = \int_0^\infty e^{-st} f(t) dt$$

with respect to s can be obtained by differentiating under the integral sign with respect to s (proof in Ref. [A1]); thus

$$F'(s) = - \int_0^\infty e^{-st} [tf(t)] dt.$$

Consequently, if $\mathcal{L}(f) = F(s)$, then

$$(1) \quad \mathcal{L}\{tf(t)\} = -F'(s); \quad .$$

differentiation of the transform of a function corresponds to the multiplication of the function by $-t$.

This property of the Laplace transformation enables us to obtain new transforms from given ones.

Example 1. We shall derive the following three formulas:

	$\mathcal{L}(f)$	$f(t)$
(2)	$\frac{1}{(s^2 + \beta^2)^2}$	$\frac{1}{2\beta^3}(\sin \beta t - \beta t \cos \beta t)$
(3)	$\frac{s}{(s^2 + \beta^2)^2}$	$\frac{t}{2\beta} \sin \beta t$
(4)	$\frac{s^2}{(s^2 + \beta^2)^2}$	$\frac{1}{2\beta}(\sin \beta t + \beta t \cos \beta t)$

From (1) and formula 8 (with $\omega = \beta$) in Table 16, Sec. 4.1, we obtain

$$\mathcal{L}(t \sin \beta t) = \frac{2\beta s}{(s^2 + \beta^2)^2}.$$

By dividing by 2β we obtain (3). From (1) and formula 9 (with $\omega = \beta$) in Table 16, Sec. 4.1, we find

$$(5) \quad \mathcal{L}(t \cos \beta t) = -\frac{(s^2 + \beta^2) - 2s^2}{(s^2 + \beta^2)^2} = \frac{s^2 - \beta^2}{(s^2 + \beta^2)^2} = \frac{1}{s^2 + \beta^2} - \frac{2\beta^2}{(s^2 + \beta^2)^2}$$

where the last equality may be readily verified by direct calculation. Hence

$$t \cos \beta t = \mathcal{L}^{-1}\left(\frac{1}{s^2 + \beta^2} - \frac{2\beta^2}{(s^2 + \beta^2)^2}\right) = \frac{1}{\beta} \sin \beta t - 2\beta^2 \mathcal{L}^{-1}\left(\frac{1}{(s^2 + \beta^2)^2}\right),$$

and we see that the last term on the right equals $f(t)$ in (2). Finally, by subtracting and adding β^2 in the numerator we have

$$\frac{s^2}{(s^2 + \beta^2)^2} = \frac{s^2 - \beta^2}{(s^2 + \beta^2)^2} + \frac{\beta^2}{(s^2 + \beta^2)^2}.$$

From (5) and (2) we see that the two terms on the right are the inverse transforms of $t \cos \beta t$ and $(\sin \beta t - t \cos \beta t)/2\beta$, respectively. This yields (4).

Similarly, if $f(t)$ satisfies the conditions of the existence theorem in Sec. 4.1 and the limit of $f(t)/t$, as t approaches 0 from the right, exists, then

$$(6) \quad \mathcal{L}\left\{\frac{f(t)}{t}\right\} = \int_s^\infty F(\tilde{s}) d\tilde{s} \quad (s > \alpha);$$

in this manner, *integration of the transform of a function $f(t)$ corresponds to the division of $f(t)$ by t .*

In fact, from the definition it follows that

$$\int_s^\infty F(\tilde{s}) d\tilde{s} = \int_s^\infty \left[\int_0^\infty e^{-\tilde{s}t} f(t) dt \right] d\tilde{s},$$

and it can be shown (cf. Ref. [A1]) that under the above assumptions the order of integration can be reversed, that is,

$$\int_s^\infty F(\tilde{s}) d\tilde{s} = \int_0^\infty \left[\int_s^\infty e^{-st} f(t) d\tilde{s} \right] dt = \int_0^\infty f(t) \left[\int_s^\infty e^{-st} d\tilde{s} \right] dt.$$

The integral over \tilde{s} on the right equals e^{-st}/t when $s > \alpha$, and, therefore,

$$\int_s^\infty F(\tilde{s}) d\tilde{s} = \int_0^\infty e^{-st} \frac{f(t)}{t} dt = \mathcal{L}\left\{\frac{f(t)}{t}\right\} \quad (s > \alpha).$$

Example 2. Find the inverse of $\ln\left(1 + \frac{\omega^2}{s^2}\right)$.

By differentiation we first have

$$-\frac{d}{ds} \ln\left(1 + \frac{\omega^2}{s^2}\right) = \frac{2\omega^2}{s(s^2 + \omega^2)} \equiv F(s).$$

Writing $F(s)$ in terms of partial fractions, we obtain

$$f(t) = \mathcal{L}^{-1}(F) = \mathcal{L}^{-1}\left\{\frac{2}{s} - 2\frac{s}{s^2 + \omega^2}\right\} = 2 - 2 \cos \omega t.$$

This function satisfies the conditions under which (6) holds. Therefore,

$$\mathcal{L}^{-1}\left\{\ln\left(1 + \frac{\omega^2}{s^2}\right)\right\} = \int_s^\infty F(\tilde{s}) d\tilde{s} = \frac{f(t)}{t}.$$

Our result is

$$\mathcal{L}^{-1}\left\{\ln\left(1 + \frac{\omega^2}{s^2}\right)\right\} = \frac{2}{t}(1 - \cos \omega t).$$

Problems

Using (1), find the Laplace transform of

- | | | |
|-----------------------|-------------------|----------------------|
| 1. te^t | 2. t^2e^{2t} | 3. $t \sin 3t$ |
| 4. t^2e^{-t} | 5. $t^2 \cos t$ | 6. $te^{-t} \cos t$ |
| 7. $te^{-t} \cosh 2t$ | 8. $t^2 \sinh 2t$ | 9. $te^{-2t} \sin t$ |

Using (6), show that

$$10. \mathcal{L}^{-1}\left\{\ln \frac{s+b}{s+a}\right\} = \frac{e^{-at} - e^{-bt}}{t} \quad 11. \mathcal{L}\left\{\frac{\sin \omega t}{t}\right\} = \text{arc cot } \frac{s}{\omega}$$

Using (6), find $f(t)$, if $\mathcal{L}(f)$ equals

- | | |
|---------------------------------|-----------------------------|
| 12. $\frac{1}{(s-4)^2}$ | 13. $\frac{2s}{(s^2+1)^2}$ |
| 14. $\ln \frac{s^2+1}{(s-1)^2}$ | 15. $\text{arc cot } (s+1)$ |

4.7 Unit Step Function

The unit step function $u_a(t)$ is defined as follows (Fig. 85):

$$(1) \quad u_a(t) = \begin{cases} 0 & \text{when } t < a \\ 1 & \text{when } t > a \end{cases} \quad (a \geq 0).$$

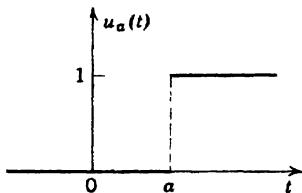


Fig. 85. Unit step function $u_a(t)$.

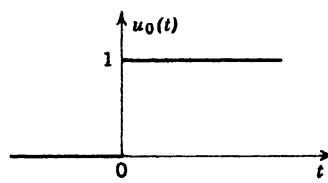


Fig. 86. Unit step function $u_0(t)$.

In particular when $a = 0$ (Fig. 86),

$$(2) \quad u_0(t) = \begin{cases} 0 & \text{when } t < 0 \\ 1 & \text{when } t > 0. \end{cases}$$

The transform of $u_a(t)$ is

$$\mathcal{L}\{u_a(t)\} = \int_0^{\infty} e^{-st} u_a(t) dt = \int_0^a e^{-st} 0 dt + \int_a^{\infty} e^{-st} 1 dt = -\frac{1}{s} e^{-st} \Big|_a^{\infty},$$

that is, assuming $s > 0$,

$$(3) \quad \mathcal{L}\{u_a(t)\} = \frac{e^{-as}}{s}.$$

The unit step function is very important. It can be considered the basic building block of certain functions whose knowledge greatly increases the use of the Laplace transformation.

Example 1. Find the transform of the function $f(t)$ shown in Fig. 87. Obviously

$$f(t) = k[u_a(t) - u_b(t)],$$

and from (3) we obtain the result

$$\mathcal{L}(f) = \frac{k}{s} (e^{-as} - e^{-bs}).$$

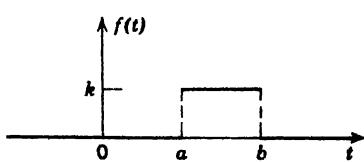


Fig. 87. $f(t)$ in Example 1.

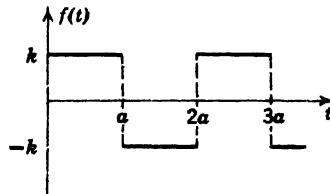


Fig. 88. Example 2.

Example 2 (Periodic square wave). Represent the square wave function $f(t)$, shown in Fig. 88, in terms of unit step functions and find its transform. Obviously,

$$f(t) = k[u_0(t) - 2u_a(t) + 2u_{2a}(t) - 2u_{3a}(t) + \dots].$$

Assuming that termwise integration is permissible and using (3), we obtain

$$\begin{aligned}\mathcal{L}(f) &= k \left[\frac{1}{s} - 2 \frac{e^{-as}}{s} + 2 \frac{e^{-2as}}{s} - + \dots \right] \\ &= \frac{k}{s} [1 - 2e^{-as}(1 - e^{-as} + e^{-2as} - + \dots)].\end{aligned}$$

The expression in parentheses (\dots) is a geometric series; using the well-known formula

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + \dots = \frac{1}{1-x} \quad |x| < 1$$

and setting $x = e^{-as}$, we find

$$\begin{aligned}\mathcal{L}(f) &= \frac{k}{s} \left(1 - \frac{2e^{-as}}{1 + e^{-as}} \right) = \frac{k}{s} \left(\frac{1 - e^{-as}}{1 + e^{-as}} \right) \\ &= \frac{k}{s} \frac{e^{-as/2}(e^{as/2} - e^{-as/2})}{e^{-as/2}(e^{as/2} + e^{-as/2})} = \frac{k}{s} \tanh \frac{as}{2} \\ &\quad (a > 0, s > 0).\end{aligned}$$

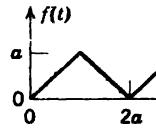


Fig. 89. Example 3.

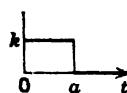
Example 3. Find the transform of the function $g(t)$ shown in Fig. 89. We see that $g(t)$ is the integral of the function $f(t)$ considered in Ex. 2, where $k = 1$. Hence, by Theorem 3 in Sec. 4.2,

$$\mathcal{L}(g) = \frac{1}{s} \mathcal{L}(f) = \frac{1}{s^2} \tanh \frac{as}{2}.$$

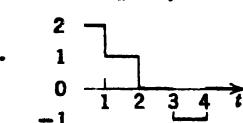
Problems

Represent the following functions in terms of unit step functions and find their Laplace transforms.

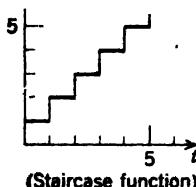
1.



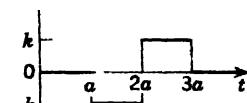
3.



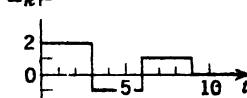
5.



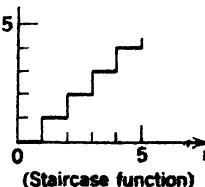
2.

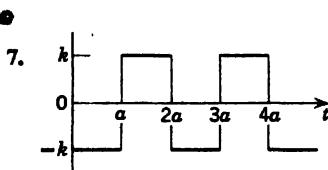
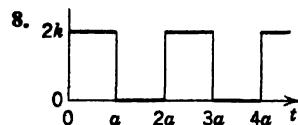
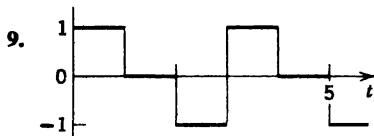
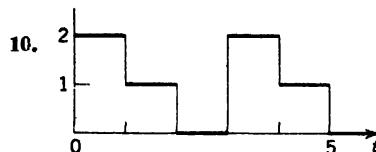


4.



6.



(Periodic, period $2a$)(Periodic, period $2a$)(Periodic, period 3)(Periodic, period 3)

11. Check the answer to Prob. 7 by using the result in Ex. 2.

12. Obtain the answer to Prob. 10 from the answer to Prob. 9.

Find and graph $f(t)$, if $\mathcal{L}(f)$ equals

13. e^{-2s}/s

16. $(1 - e^{-2s})/s$

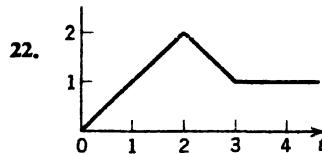
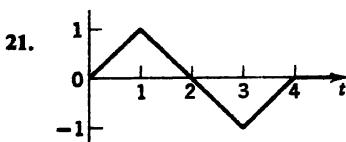
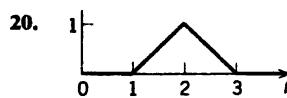
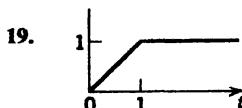
14. $(1 - e^{-s})/s$

17. e^{-s}/s^2

15. $(e^{-3s} - e^{-s})/s$

18. $(2e^{-as} - 1)/s$

Applying Theorem 3 in Sec. 4.2 to suitable step functions, find the Laplace transforms of the following functions.



23. Check the answer to Prob. 19 by using the definition of the Laplace transform and integrating.

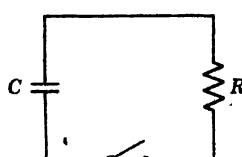


Fig. 90. Problem 24.

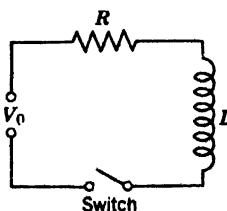


Fig. 91. Problem 25.

24. A capacitor of capacitance C is charged so that its potential is V_0 . At $t = 0$ the switch in Fig. 90 is closed and the capacitor starts discharging through the resistor of resistance R . Using the Laplace transform find the charge $q(t)$ on the capacitor.
25. Find the current $i(t)$ in the circuit in Fig. 91, assuming that no current flows when $t \leq 0$, and the switch is closed at $t = 0$.

4.8 Shifting on the t -Axis

In Sec. 4.1 it was shown that

$$\text{if } F(s) = \mathcal{L}\{f(t)\}, \quad \text{then} \quad F(s - a) = \mathcal{L}\{e^{at}f(t)\}.$$

This theorem answers the question how the function changes when the argument s of the *transform* is replaced by $s - a$ (*shifting on the s -axis*). Conversely, we may ask how the *transform* of a function $f(t)$ changes when the argument t of the *function* is replaced by $t - a$ (*shifting on the t -axis*).

The latter question is of importance in connection with systems which are acted upon by a disturbance beginning not at $t = 0$ but at some later time $t = a$.

Suppose that the transform

$$F(s) = \mathcal{L}\{f(\tau)\} = \int_0^\infty e^{-st}f(\tau) d\tau$$

of a given function $f(\tau)$ exists. Multiplying by e^{-as} , where a is some positive constant, we have

$$e^{-as}F(s) = \int_0^\infty e^{-s(\tau+a)}f(\tau) d\tau.$$

Substituting $\tau + a = t$ in the integral, we obtain

$$e^{-as}F(s) = \int_a^\infty e^{-st}f(t-a) dt.$$

We want to write this integral as an integral from 0 to ∞ . For this purpose we may replace $f(t-a)$ by the function which is zero on the interval $0 \leq t < a$ and is equal to $f(t-a)$ when $t > a$. From the definition of the unit step function it follows that the function

$$f(t-a)u_a(t) = \begin{cases} 0 & (t < a) \\ f(t-a) & (t > a) \end{cases}$$

has the desired property. (Figure 92 shows an illustrative example.) Therefore,

$$e^{-as}F(s) = \int_a^\infty e^{-st}f(t-a)u_a(t) dt = \mathcal{L}\{f(t-a)u_a(t)\}.$$

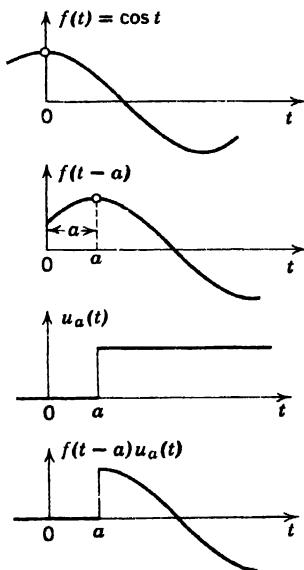


Fig. 92. $f(t-a)u_a(t)$ where $f(t) = \cos t$.

We may sum up our result as follows.

Theorem 1 (Second shifting theorem). If

$$F(s) = \mathcal{L}\{f(t)\},$$

then for any positive constant a ,

$$(1) \quad e^{-as}F(s) = \mathcal{L}\{f(t-a)u_a(t)\}$$

where the unit step function $u_a(t)$ is defined by (1), Sec. 4.7.

Example 1. Find the inverse of e^{-3s}/s^3 .

Since $\mathcal{L}^{-1}(1/s^3) = t^2/2$ (cf. Table 16 in Sec. 4.1), Theorem 1 gives

$$\mathcal{L}^{-1}(e^{-3s}/s^3) = \frac{1}{2}(t-3)^2 u_3(t) \quad (\text{cf. Fig. 93}).$$

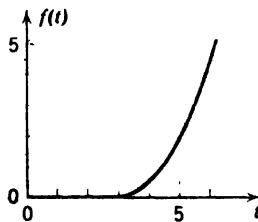


Fig. 93. Example 1.

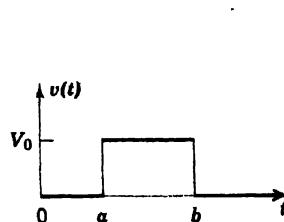
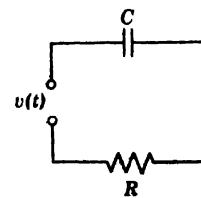


Fig. 94. Example 2.



Example 2 (Response of an RC-circuit to a single square wave). Find the current $i(t)$ flowing in the circuit in Fig. 94, if a single square wave with voltage of height V_0 is applied. The circuit is assumed to be quiescent before the square wave is applied.

The equation of the circuit is (cf. Sec. 1.9)

$$Ri(t) + \frac{q(t)}{C} = Ri(t) + \frac{1}{C} \int_0^t i(\tau) d\tau = v(t)$$

where $v(t)$ can be represented in terms of two unit step functions:

$$v(t) = V_0[u_a(t) - u_b(t)].$$

Using Theorem 3 in Sec. 4.2 and formula (3) in Sec. 4.7, we obtain the subsidiary equation

$$RI(s) + \frac{I(s)}{sC} = \frac{V_0}{s} [e^{-as} - e^{-bs}].$$

The solution of this equation may be written

$$I(s) = F(s)(e^{-as} - e^{-bs}) \quad \text{where} \quad F(s) = \frac{V_0/R}{s + 1/RC}.$$

From Table 16 in Sec. 4.1 we have

$$\mathcal{L}^{-1}(F) = \frac{V_0}{R} e^{-t/RC}.$$

Hence, by the use of Theorem 1, we obtain the solution (Fig. 95)

$$\begin{aligned} i(t) &= \mathcal{L}^{-1}(I) = \mathcal{L}^{-1}\{e^{-as}F(s)\} - \mathcal{L}^{-1}\{e^{-bs}F(s)\} \\ &= \frac{V_0}{R}[e^{-(t-a)/RC}u_a(t) - e^{-(t-b)/RC}u_b(t)]; \end{aligned}$$

that is, $i = 0$ when $t < a$, and

$$i(t) = \begin{cases} K_1 e^{-t/RC} & \text{when } a < t < b \\ (K_1 - K_2)e^{-t/RC} & \text{when } t > b \end{cases} \quad \text{where} \quad \begin{aligned} K_1 &= \frac{V_0}{R} e^{a/RC} \\ K_2 &= \frac{V_0}{R} e^{b/RC}. \end{aligned}$$

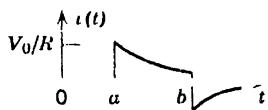
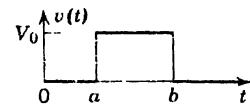


Fig. 95. Voltage and current in Example 2.

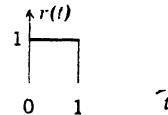


Fig. 96. Example 3.

Example 3 (Response of a damped vibrating system to a single square wave). Determine the response of the damped vibrating system corresponding to the equation

$$y'' + 3y' + 2y = r(t)$$

where $r(t) = 1$ when $0 < t < 1$ and 0 otherwise (Fig. 96); assume that $y(0) = 0$ and $y'(0) = 0$.

We have

$$r(t) = u_0(t) - u_1(t).$$

The subsidiary equation is

$$s^2 Y + 3s Y + 2Y = \frac{1}{s}(1 - e^{-s}).$$

Solving for Y , we obtain

$$Y(s) = F(s)(1 - e^{-s}) \quad \text{where} \quad F(s) = \frac{1}{s(s+1)(s+2)}.$$

Using the method of partial fractions [cf. Sec. 4.4], we find

$$f(t) = \mathcal{L}^{-1}(F) = \frac{1}{2} - e^{-t} + \frac{1}{2}e^{-2t}.$$

Therefore, by Theorem 1, we have

$$\mathcal{L}^{-1}\{e^{-s}F(s)\} = f(t-1)u_1(t) = \begin{cases} 0 & (t < 1) \\ \frac{1}{2} - e^{-(t-1)} + \frac{1}{2}e^{-2(t-1)} & (t > 1) \end{cases}$$

This yields the solution (cf. Fig. 97 on the next page)

$$y(t) = \mathcal{L}^{-1}(Y) = f(t) - f(t-1)u_1(t) = \begin{cases} \frac{1}{2} - e^{-t} + \frac{1}{2}e^{-2t} & (0 \leq t < 1) \\ K_1 e^{-t} - K_2 e^{-2t} & (t \geq 1) \end{cases}$$

where $K_1 = e - 1$ and $K_2 = (e^2 - 1)/2$.

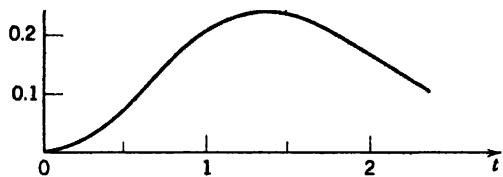


Fig. 97. Output in Example 3.

Problems

Graph the following functions and find their Laplace transforms.

1. $(t - 1)u_1(t)$
2. $(t - 2)u_2(t)$
3. $tu_1(t)$
4. $e^t u_1(t)$
5. $t^2 u_1(t)$
6. $e^t u_2(t)$
7. $e^{kt} u_a(t)$
8. $(t - 4)^3 u_4(t)$
9. $e^{-2t} u_x(t)$
10. $u_{2\pi/\omega}(t) \sin \omega t$
11. $u_{2\pi/\omega}(t) \sin \omega t$
12. $u_x(t) \cos t$
13. $K \sin \omega t$ when $0 < t < \pi/\omega$ and 0 otherwise
14. $K \sin t$ when $0 < t < 2\pi$ and 0 otherwise
15. $K \cos \omega t$ when $0 < t < \pi/\omega$ and 0 otherwise
16. $K \sin t$ when $2\pi < t < 4\pi$ and 0 otherwise
17. Kt when $0 < t < a$ and 0 otherwise
18. Kt when $t > a$ and 0 otherwise
19. Solve Prob. 16 by applying Theorem 1 to the answer of Prob. 14.

Find and graph the inverse Laplace transforms of the following functions.

20. e^{-2s}/s
21. e^{-s}/s^3
22. $e^{-2s}/(s - 2)$
23. $e^{-s}/(s^2 + \pi^2)$
24. $e^{-s}/(s - 3)$
25. $(1 - e^{-\pi s})/(s^2 + 4)$

Find the current $i(t)$ in the LC-circuit in Fig. 98, assuming $L = 1$ henry, $C = 1$ farad, zero initial current and charge on the capacitor, and $v(t)$ as follows.

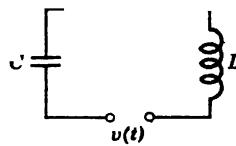


Fig. 98. LC-circuit.

26. $v = 1$ when $0 < t < 1$ and 0 otherwise
27. $v = t$ when $0 < t < a$ and 0 otherwise
28. $v = t$ when $0 < t < a$ and $v = a$ when $t > a$
29. $v = 1 - e^{-t}$ when $0 < t < \pi$ and 0 otherwise
30. Show that if $\mathcal{L}\{f(t)\} = F(s)$, then $\mathcal{L}\{f(at)\} = F(s/a)/a$. Using this result, derive the transform of $\sin \omega t$ from that of $\sin t$.

4.9 Periodic Functions

The Laplace transforms of some special periodic functions were already considered in Sec. 4.7, and we shall now approach the problem of determining such transforms in a systematic way.

Let $f(t)$ be a function which is defined for all positive t and has the period $p (> 0)$, that is,

$$f(t + p) = f(t) \quad \text{for all } t > 0.$$

If $f(t)$ is piecewise continuous over an interval of length p , then its Laplace transform exists, and we can write the integral from zero to infinity as the series of integrals over successive periods:

$$\mathcal{L}(f) = \int_0^\infty e^{-st} f(t) dt = \int_0^p e^{-st} f(t) dt + \int_p^{2p} e^{-st} f(t) dt + \int_{2p}^{3p} e^{-st} f(t) dt + \dots$$

If we substitute $t = \tau + p$ in the second integral, $t = \tau + 2p$ in the third integral, \dots , $t = \tau + (n - 1)p$ in the n th integral, \dots , then the new limits are 0 and p . Since

$$f(\tau + p) = f(\tau), \quad f(\tau + 2p) = f(\tau),$$

etc., we thus obtain

$$\mathcal{L}(f) = \int_0^p e^{-s\tau} f(\tau) d\tau + \int_0^p e^{-s(\tau+p)} f(\tau) d\tau + \int_0^p e^{-s(\tau+2p)} f(\tau) d\tau + \dots$$

Taking the factors which do not depend on τ out from under the integral signs, we find

$$\mathcal{L}(f) = [1 + e^{-sp} + e^{-2sp} + \dots] \int_0^p e^{-s\tau} f(\tau) d\tau.$$

The series in brackets $[\dots]$ is a geometric series whose sum is $1/(1 - e^{-sp})$ (cf. also Ex. 2 in Sec. 4.7). The following result is therefore established.

Theorem 1 (Transform of periodic functions). *The Laplace transform of a piecewise continuous periodic function $f(t)$ with period p is*

$$(1) \quad \mathcal{L}(f) = \frac{1}{1 - e^{-sp}} \int_0^p e^{-st} f(t) dt \quad (s > 0).$$

Example 1 (Half-wave rectifier). Find the Laplace transform of the following function $f(t)$ with period $p = 2\pi/\omega$:

$$f(t) = \begin{cases} \sin \omega t & \text{when } 0 < t < \pi/\omega, \\ 0 & \text{when } \pi/\omega < t < 2\pi/\omega. \end{cases}$$

Note that this function is the half-wave rectification of $\sin \omega t$ (Fig. 99). From (1) we obtain

$$\mathcal{L}(f) = \frac{1}{1 - e^{-2\pi s/\omega}} \int_0^{\pi/\omega} e^{-st} \sin \omega t dt.$$

Fig. 99. Half-wave rectification of $\sin \omega t$.

Integrating by parts or noting that the integral is the imaginary part of the integral

$$\int_0^{\pi/\omega} e^{(-s+i\omega)t} dt = \frac{1}{-s + i\omega} e^{(-s+i\omega)t} \Big|_0^{\pi/\omega} = \frac{-s - i\omega}{s^2 + \omega^2} (-e^{-s\pi/\omega} - 1)$$

we obtain the result

$$\mathcal{L}(f) = \frac{\omega(1 + e^{-\pi s/\omega})}{(s^2 + \omega^2)(1 - e^{-2\pi s/\omega})} = \frac{\omega}{(s^2 + \omega^2)(1 - e^{-\pi s/\omega})}.$$

Example 2 (Saw-tooth wave). Find the Laplace transform of the function (Fig. 100)

$$f(t) = \frac{k}{p} t \quad \text{when } 0 < t < p, \quad f(t+p) = f(t).$$

Integration by parts yields

$$\begin{aligned} \int_0^p e^{-st} t dt &= -\frac{t}{s} e^{-st} \Big|_0^p + \frac{1}{s} \int_0^p e^{-st} dt \\ &= -\frac{p}{s} e^{-sp} - \frac{1}{s^2} (e^{-sp} - 1), \end{aligned}$$

and from (1) we thus obtain the result

$$\mathcal{L}(f) = \frac{k}{ps^2} - \frac{ke^{-ps}}{s(1 - e^{-ps})} \quad (s > 0).$$

Example 3 (Staircase function). Find the Laplace transform of the staircase function (Fig. 101)

$$g(t) = kn \quad [np < t < (n+1)p, \quad n = 0, 1, 2, \dots].$$

Since $g(t)$ is the difference of the functions $h(t) = kt/p$ (whose transform is $k/p s^2$) and $f(t)$ in Ex. 2, we obtain

$$\mathcal{L}(g) = \mathcal{L}(h) - \mathcal{L}(f) = \frac{ke^{-ps}}{s(1 - e^{-ps})} \quad (s > 0).$$

The next example illustrates the application of the Laplace transformation to a linear nonhomogeneous differential equation which has a periodic function on the right-hand side.



Fig. 100. Saw-tooth wave.

Fig. 101. Staircase function.

Example 4 (Forced oscillations). Solve the differential equation

$$(2) \quad y'' + 2y' + 10y = r(t)$$

where (Fig. 102)

$$r(t) = \begin{cases} 1 & (0 < t < \pi) \\ -1 & (\pi < t < 2\pi) \end{cases} \quad r(t+2\pi) = r(t).$$

We have (cf. Ex. 2 in Sec. 4.7)

$$R(s) = \mathcal{L}(r) = \frac{1}{s} [1 - 2e^{-\pi s} + 2e^{-2\pi s} - 2e^{-3\pi s} + \dots].$$

The subsidiary equation of (2) is

$$s^2 Y - sy(0) - y'(0) + 2[sY - y(0)] + 10Y = R(s).$$

By solving for Y we obtain

$$(3) \quad Y(s) = \frac{(s+2)y(0) + y'(0)}{s^2 + 2s + 10} + \frac{R(s)}{s^2 + 2s + 10}.$$

The roots of the denominator are

$$a = -1 + 3i \quad \text{and} \quad \bar{a} = -1 - 3i.$$

The inverse transform of the first term in (3) is

$$(4) \quad \begin{aligned} \mathcal{L}^{-1}\left(\frac{(s+2)y(0) + y'(0)}{s^2 + 2s + 10}\right) &= \mathcal{L}^{-1}\left(\frac{(s+1)y(0)}{(s+1)^2 + 9} + \frac{y(0) + y'(0)}{(s+1)^2 + 9}\right) \\ &= e^{-t}\left(y(0) \cos 3t + \frac{y(0) + y'(0)}{3} \sin 3t\right). \end{aligned}$$

This is the solution of the homogeneous equation corresponding to (2). The last term in (3) is

$$\frac{R(s)}{s^2 + 2s + 10} = \frac{1}{s[(s+1)^2 + 9]} \{1 - 2e^{-\pi s} + \dots\}.$$

Now we know that

$$\mathcal{L}^{-1}\left(\frac{1}{(s+2)^2 + 9}\right) = \frac{1}{3} e^{-t} \sin 3t.$$

Therefore, according to Theorem 3 in Sec. 4.2,

$$(5) \quad \mathcal{L}^{-1}\left(\frac{1}{s((s+1)^2 + 9)}\right) = \frac{1}{3} \int_0^t e^{-\tau} \sin 3\tau d\tau = \frac{1}{10} [1 - h(t)]$$

where

$$(6) \quad h(t) = e^{-t}(\cos 3t + \frac{1}{3} \sin 3t).$$

Denoting the function on the right-hand side of (5) by $f(t)$, that is,

$$f(t) = \frac{1}{10} - \frac{h(t)}{10},$$

and applying Theorem 1, Sec. 4.8, we thus obtain

$$(7) \quad \mathcal{L}^{-1}\left(\frac{R(s)}{s^2 + 2s + 10}\right) = f(t) - 2f(t-\pi)u_\pi(t) + 2f(t-2\pi)u_{2\pi}(t) - \dots$$

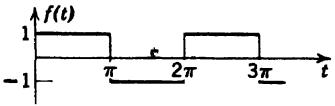


Fig. 102. Input $r(t)$ in Example 4.

Now in the terms on the right,

$$h(t - \pi) = e^{-(t-\pi)}[\cos 3(t - \pi) + \frac{1}{3} \sin 3(t - \pi)] = -h(t)e^{\pi}$$

and similarly,

$$h(t - 2\pi) = h(t)e^{2\pi}, \quad h(t - 3\pi) = -h(t)e^{3\pi},$$

etc. It follows that in (7),

$$f(t - \pi) = \frac{1}{10} - \frac{h(t - \pi)}{10} = \frac{1}{10} + \frac{h(t)}{10}e^{\pi}$$

$$f(t - 2\pi) = \frac{1}{10} - \frac{h(t - 2\pi)}{10} = \frac{1}{10} - \frac{h(t)}{10}e^{2\pi}$$

etc., and the right-hand side of (7) equals

$$\frac{1}{10} - \frac{h(t)}{10} \quad \text{when } 0 < t < \pi,$$

$$- \frac{1}{10} - \frac{h(t)}{10} - \frac{h(t)}{5}e^{\pi} \quad \text{when } \pi < t < 2\pi,$$

$$\frac{1}{10} - \frac{h(t)}{10} - \frac{h(t)}{5}(e^{\pi} + e^{2\pi}) \quad \text{when } 2\pi < t < 3\pi,$$

$$\frac{(-1)^n}{10} + \frac{h(t)}{10} - \frac{h(t)}{5}(1 + e^{\pi} + \cdots + e^{n\pi}) \quad \text{when } n\pi < t < (n+1)\pi.$$

Summing the finite geometric progression in parentheses, we see that the expression in the last line may be written

$$\frac{(-1)^n}{10} + \frac{h(t)}{10} - \frac{h(t)}{5} \left(\frac{e^{(n+1)\pi} - 1}{e^{\pi} - 1} \right).$$

If we write $h(t)$ at length, this becomes

$$(8) \quad \begin{aligned} & \left[\frac{1}{10} + \frac{1}{5(e^{\pi} - 1)} \right] e^{-t} \left(\cos 3t + \frac{1}{3} \sin 3t \right) \\ & + \frac{(-1)^n}{10} - \frac{1}{5(e^{\pi} - 1)} e^{-|t-(n+1)\pi|} \left(\cos 3t + \frac{1}{3} \sin 3t \right), \end{aligned}$$

and the solution of (2) on the interval $n\pi < t < (n+1)\pi$ is the sum of (8) and the right-hand side of (4).

Clearly the function in (4) and the function in the first line of (8) approach zero as t approaches infinity. The function in the last line of (8) is a periodic function of t with period 2π and thus represents the *steady-state solution* of our problem. To see this more distinctly we denote that function by $y_2(t)$ and set $t - n\pi = \tau$. Then $t = \tau + n\pi$, and τ ranges from 0 to π as t ranges from $n\pi$ to $(n+1)\pi$. Also

$$\cos 3t = \cos(3\tau + 3n\pi) = \cos 3\tau \cos 3n\pi = (-1)^n \cos 3\tau,$$

$$\sin 3t = \sin(3\tau + 3n\pi) = \sin 3\tau \cos 3n\pi = (-1)^n \sin 3\tau,$$

and the function y_2 becomes

$$(9) \quad \begin{aligned} y_2 &= (-1)^n \left[\frac{1}{10} - \frac{e^{-\tau+\pi}}{5(e^{\pi} - 1)} \left(\cos 3\tau + \frac{1}{3} \sin 3\tau \right) \right] \\ &\approx (-1)^n [0.1 - 0.22e^{-\tau} \cos(3\tau - 0.32)] \end{aligned}$$

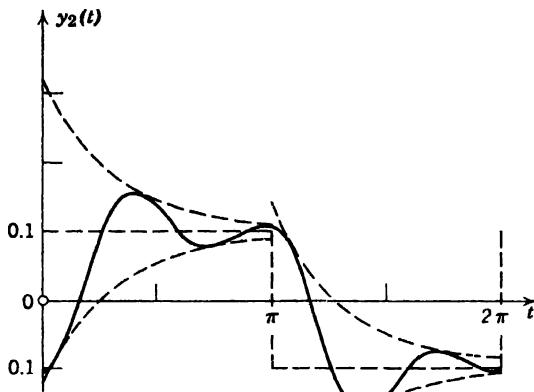


Fig. 103. Steady-state output (9) in Example 4.

where $0 < \tau < \pi$ and $t = \tau + n\pi$, $n = 0, 1, \dots$. Figure 103 shows this function; the graph for $0 < t < \pi$ is obtained from (9) with $n = 0$ and $t = \tau$, and for $\pi < t < 2\pi$ it is obtained from (9) with $n = 1$ and $t = \tau + \pi$.

We have the surprising result that the output y_2 tends to oscillate more rapidly than the input $r(t)$. The reason for this unexpected behavior of our system will become obvious from our later consideration of Fourier series in Sec. 8.7.

Problems

1. Apply Theorem 1 to the function $f(t) = 1$ which is periodic with any period p .
2. Show that the solution y_2 in Ex. 4 is continuous at 0 and π .
3. Find the Laplace transform of the half-wave rectification of $-\sin \omega t$ (Fig. 104).

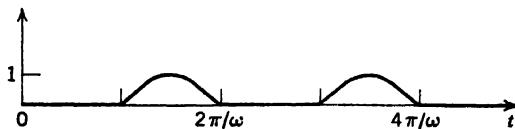


Fig. 104. Problem 3.

4. Solve Prob. 3 by applying Theorem 1, Sec. 4.8, to the result of Ex. 1.
5. Find the Laplace transform of the full-wave rectification of $\sin \omega t$ (Fig. 105).

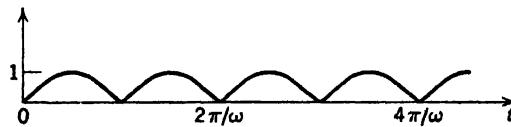


Fig. 105. Problem 5.

6. Check the answer to Prob. 5 by using the results of Ex. 1 and Prob. 3.
7. Find the Laplace transform of $|\cos \omega t|$.

Graph the following functions which are assumed to have the period 2π and find their transforms.

8. $f(t) = 2\pi - t \quad (0 < t < 2\pi)$
9. $f(t) = e^t \quad (0 < t < 2\pi)$
10. $f(t) = t^2 \quad (0 < t < 2\pi)$
11. $f(t) = K \sin(t/2) \quad (0 < t < 2\pi)$
12. $f(t) = 1$ when $0 < t < \pi$, $f(t) = -1$ when $\pi < t < 2\pi$
13. $f(t) = t$ when $0 < t < \pi$, $f(t) = \pi - t$ when $\pi < t < 2\pi$
14. $f(t) = 0$ when $0 < t < \pi$, $f(t) = t - \pi$ when $\pi < t < 2\pi$
15. $f(t) = t$ when $0 < t < \pi$, $f(t) = 0$ when $\pi < t < 2\pi$

16. Check the answer to Prob. 13 by subtracting the result of Prob. 14 from that of Prob. 15.

17. Solve Prob. 14 by applying Theorem 1, Sec. 4.8, to the result of Prob. 15.

18. A capacitor ($C = 1$ farad) is charged to the potential $V_0 = 100$ volts and discharged starting at $t = 0$ by closing the switch in Fig. 106. Find the current in the circuit and the charge in the capacitor.

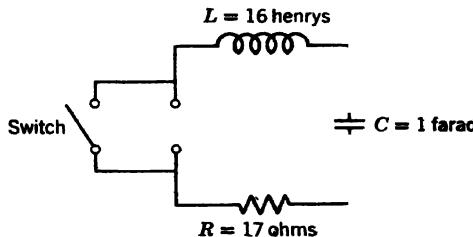


Fig. 106. Problem 18.

19. Find the ramp-wave response of the RC -circuit in Fig. 107 assuming that the circuit is quiescent at $t = 0$.

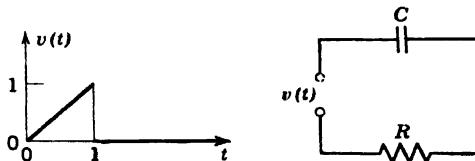


Fig. 107. Problem 19.

20. Solve Prob. 19 by means of the Laplace transformation, starting from the equation of the form $RI' + (1/C)i = v'$.

21. Solve Prob. 19 without the use of the Laplace transformation. Explain the reason for the jump of $i(t)$ at $t = 1$.

22. Find the current $i(t)$ in the RC -circuit in Fig. 107 when $v(t) = \sin \omega t$ ($0 < t < \pi/\omega$), $v(t) = 0$ ($t > \pi/\omega$), and $i(0) = 0$.

23. Find the steady-state current in the circuit in Fig. 108.

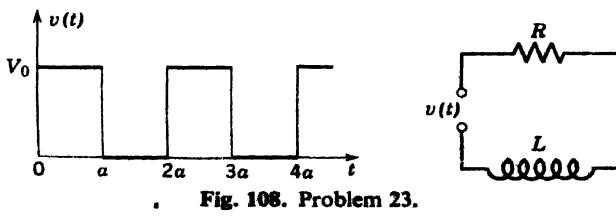


Fig. 108. Problem 23.

24. Solve Prob. 23 without the use of the Laplace transformation.
25. Using the Laplace transformation, show that the current $i(t)$ in the RLC-circuit in Fig. 109 (constant electromotive force V_0 , zero initial current and charge) is

$$i(t) = \begin{cases} (K/\omega^2)e^{-\alpha t} \sin \omega^* t & \text{if } \omega^{*2} > 0 \\ Kte^{-\alpha t} & \text{if } \omega^{*2} = 0 \\ (K/\beta)e^{-\alpha t} \sinh \beta t & \text{if } \omega^{*2} = -\beta^2 < 0; \end{cases}$$

here $K = V_0/L$, $\alpha = R/2L$, $\omega^{*2} = (1/LC) - \alpha^2$.

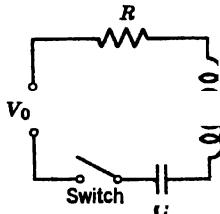


Fig. 109. Problem 25.

26. Find the current in the circuit in Prob. 25, assuming that the electromotive force applied at $t = 0$ is $V_0 \sin pt$, and the current and charge at $t = 0$ are zero.
27. Find the current in Prob. 25, assuming that at $t = 0$ the charge on the capacitor is zero, $i(0) = i_0$, and $\omega^{*2} > 0$.
28. Find the current in the RLC-circuit in Fig. 109, if a battery of electromotive force V_0 is connected to the circuit at $t = 0$ and short-circuited at $t = a$. Assume that the initial current and charge are zero, and ω^{*2} , as defined in Prob. 25, is positive.
29. Find the current $i(t)$ in the circuit in Fig. 110, assuming that $i(0) = 0$.

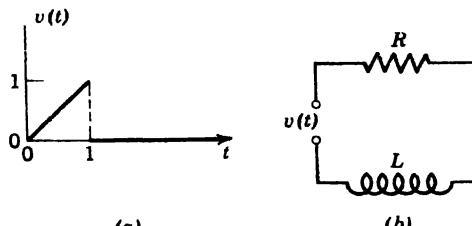


Fig. 110. Problem 29.

30. Steady current is flowing in the circuit in Fig. 111 with the switch closed. At $t = 0$ the switch is opened. Find the current $i(t)$.

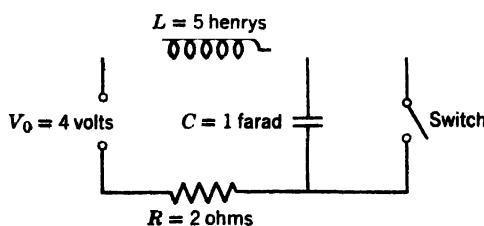


Fig. 111. Problem 30.

31. Find the steady-state current in the circuit in Fig. 110b, if $v(t) = t$ when $0 < t < 1$ and $v(t + 1) = v(t)$ as shown in Fig. 112.

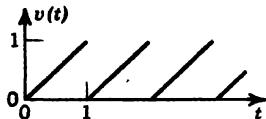


Fig. 112. Problem 31.

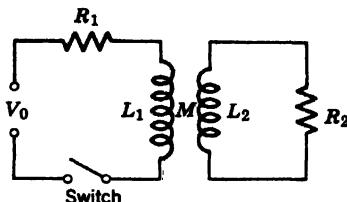


Fig. 113. Problem 32.

32. The circuits in Fig. 113 are coupled by mutual inductance M , and at $t = 0$ the currents $i_1(t)$ and $i_2(t)$ are zero. Applying Kirchhoff's second law (Sec. 1.9), show that the differential equations for the currents are

$$\begin{aligned} L_1 i_1' + R_1 i_1 &= Mi_2' + V_0 u_0(t) \\ L_2 i_2' + R_2 i_2 &= Mi_1' \end{aligned}$$

where $u_0(t)$ is the unit step function. Setting $\mathcal{L}(i_1) = I_1(s)$ and $\mathcal{L}(i_2) = I_2(s)$, show that the subsidiary equations are

$$\begin{aligned} L_1 s I_1 + R_1 I_1 &= M s I_2 + \frac{V_0}{s} \\ L_2 s I_2 + R_2 I_2 &= M s I_1. \end{aligned}$$

Assuming that $A \equiv L_1 L_2 - M^2 > 0$, show that the expression for I_2 , obtained by solving these algebraic equations, may be written

$$I_2 = \frac{K}{(s + \alpha)^2 + \omega^{*2}}$$

where $K = A^{-1} V_0 M$, $\alpha = (2A)^{-1} (R_1 L_2 + R_2 L_1)$, $\omega^{*2} = A^{-1} R_1 R_2 - \alpha^2$. Show that $i_2(t) = \mathcal{L}^{-1}(I_2)$ is of the same form as $i(t)$ in Prob. 25, where the constants K , α , and ω^* are now those defined in the present problem.

33. Find $i_1(t)$ in Prob. 32, assuming that $\omega^{*2} > 0$.
34. Show that when $A = 0$ in Prob. 32, then $i_2(t) = K_0 e^{-\alpha t}$ where $K_0 = B V_0 M$, $\alpha = B R_1 R_2$, $B = (R_1 L_2 + R_2 L_1)^{-1}$.
35. Suppose that in the first circuit in Fig. 113 the switch is closed and a steady current V_0/R_1 is flowing in it. At $t = 0$ the switch is opened. Find the secondary current $i_2(t)$. Hint: note that $i_1(0) = V_0/R_1$ and $i_1 = 0$ when $t > 0$.

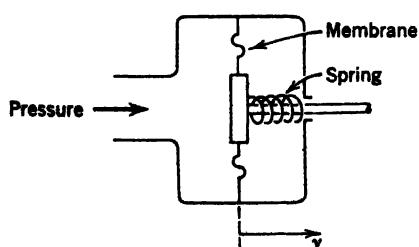
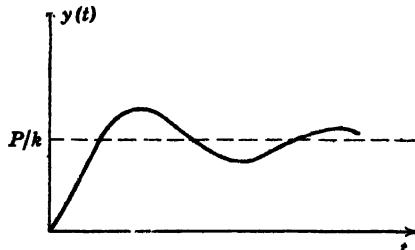


Fig. 114. Problem 36.

Fig. 115. Displacement $y(t)$ in Problem 36.

36. Figure 114 shows a system for automatic control of pressure. $y(t)$ is the displacement, where $y = 0$ corresponds to the equilibrium position due to a given constant pressure. We make the following assumptions. The damping of the system is proportional to the velocity y' . For $t < 0$, the system is at rest. At $t = 0$ the pressure is suddenly increased in the form of a unit step function. Show that the corresponding differential equation is

$$my'' + cy' + ky = Pu_0(t)$$

(m = effective mass of the moving parts, c = damping constant, k = spring modulus, P = force due to the increase of pressure at $t = 0$). Using the Laplace transformation, show that if $c^2 < 4mk$, then (Fig. 115)

$$y(t) = \frac{P}{k} [1 - e^{-\alpha t} \sqrt{1 + (\alpha/\omega^*)^2} \cos(\omega^*t + \theta)]$$

where $\alpha = c/2m$, $\omega^* = \sqrt{(k/m) - \alpha^2}$ (> 0), $\tan \theta = -\alpha/\omega^*$.

37. Solve Prob. 36 when $c^2 > 4mk$.
 38. Solve Prob. 36 when $c^2 = 4mk$.
 39. Two flywheels (moments of inertia M_1 and M_2) are connected by an elastic shaft (moment of inertia negligible), and are rotating with constant angular velocity ω . At $t = 0$ a constant retarding couple P is applied to the first wheel. Find the subsequent angular velocity $v(t)$ of the other wheel.
 40. Same conditions as in Prob. 39, but the retarding couple is applied during the interval $0 < t < 1$ only. Find $v(t)$.

4.10 Table 17.

Some Laplace Transforms

Remark. For a more extensive list of Laplace transforms and their inverse transforms, see Ref. [B11] in Appendix 1.

	$F(s) = \mathcal{L}\{f(t)\}$	$f(t)$
1	$1/s$	1
2	$1/s^2$	t
3	$1/s^n, \quad (n = 1, 2, \dots)$	$t^{n-1}/(n - 1)!$
4	$1/\sqrt{s}$	$1/\sqrt{\pi t}$
5	$1/s^{3/2}$	$2\sqrt{t/\pi}$
6	$1/s^a \quad (a > 0)$	$t^{a-1}/\Gamma(a)$
7	$\frac{1}{s-a}$	e^{at}
8	$\frac{1}{(s-a)^2}$	te^{at}
9	$\frac{1}{(s-a)^n} \quad (n = 1, 2, \dots)$	$\frac{1}{(n-1)!} t^{n-1} e^{at}$
10	$\frac{1}{(s-a)^k} \quad (k > 0)$	$\frac{1}{\Gamma(k)} t^{k-1} e^{at}$
11	$\frac{1}{(s-a)(s-b)} \quad (a \neq b)$	$\frac{1}{(a-b)} (e^{at} - e^{bt})$
12	$\frac{s}{(s-a)(s-b)} \quad (a \neq b)$	$\frac{1}{(a-b)} (ae^{at} - be^{bt})$
13	$\frac{1}{s^2 + \omega^2}$	$\frac{1}{\omega} \sin \omega t$
14	$\frac{s}{s^2 + \omega^2}$	$\cos \omega t$
15	$\frac{1}{s^2 - a^2}$	$\frac{1}{a} \sinh at$
16	$\frac{s}{s^2 - a^2}$	$\cosh at$
17	$\frac{1}{(s-a)^2 + \omega^2}$	$\frac{1}{\omega} e^{at} \sin \omega t$
18	$\frac{s-a}{(s-a)^2 + \omega^2}$	$e^{at} \cos \omega t$

	$F(s) = \mathcal{L}\{f(t)\}$	$f(t)$
19	$\frac{1}{s(s^2 + \omega^2)}$	$\frac{1}{\omega^2} (1 - \cos \omega t)$
20	$\frac{1}{s^2(s^2 + \omega^2)}$	$\frac{1}{\omega^3} (\omega t - \sin \omega t)$
21	$\frac{1}{(s^2 + \omega^2)^2}$	$\frac{1}{2\omega^3} (\sin \omega t - \omega t \cos \omega t)$
22	$\frac{s}{(s^2 + \omega^2)^2}$	$\frac{t}{2\omega} \sin \omega t$
23	$\frac{s^2}{(s^2 + \omega^2)^2}$	$\frac{1}{2\omega} (\sin \omega t + \omega t \cos \omega t)$
24	$\frac{s}{(s^2 + a^2)(s^2 + b^2)} \quad (a^2 \neq b^2)$	$\frac{1}{b^2 - a^2} (\cos at - \cos bt)$
25	$\frac{1}{s^4 + 4a^4}$	$\frac{1}{4a^3} (\sin at \cosh at - \cos at \sinh at)$
26	$\frac{s}{s^4 + 4a^4}$	$\frac{1}{2a^2} \sin at \sinh at$
27	$\frac{1}{s^4 - a^4}$	$\frac{1}{2a^3} (\sinh at - \sin at)$
28	$\frac{s}{s^4 - a^4}$	$\frac{1}{2a^2} (\cosh at - \cos at)$
29	$\sqrt{s-a} - \sqrt{s-b}$	$\frac{1}{2\sqrt{\pi t^3}} (e^{bt} - e^{at})$
30	$\frac{1}{\sqrt{s+a} \sqrt{s+b}}$	$e^{-(a+b)t/2} I_0\left(\frac{a-b}{2}t\right) \quad (\text{cf. Sec. 3.7})$
31	$\frac{1}{\sqrt{s^2 + a^2}}$	$J_0(at) \quad (\text{cf. Sec. 3.5})$
32	$\frac{s}{(s-a)^{3/2}}$	$\frac{1}{\sqrt{\pi t}} e^{at}(1 + 2at)$
33	$\frac{1}{(s^2 - a^2)^k} \quad (k > 0)$	$\frac{\sqrt{\pi}}{\Gamma(k)} \left(\frac{t}{2a}\right)^{k-1/2} I_{k-1/2}(at) \quad (\text{cf. Sec. 3.7})$

	$F(s) = \mathcal{L}\{f(t)\}$	$f(t)$
34	$\frac{1}{s} e^{-k/s}$	$J_0(2\sqrt{kt})$ (cf. Sec. 3.5)
35	$\frac{1}{\sqrt{s}} e^{-k/s}$	$\frac{1}{\sqrt{\pi t}} \cos 2\sqrt{kt}$
36	$\frac{1}{s^{3/2}} e^{k/s}$	$\frac{1}{\sqrt{\pi k}} \sinh 2\sqrt{kt}$
37	$\frac{1}{s} \ln s$	$-\ln t - \gamma$ ($\gamma \approx 0.5772$; cf. Sec. 3.7)
38	$\ln \frac{s-a}{s-b}$	$\frac{1}{t} (e^{bt} - e^{at})$
39	$\ln \frac{s^2 + \omega^2}{s^2}$	$\frac{2}{t} (1 - \cos \omega t)$
40	$\ln \frac{s^2 - a^2}{s^2}$	$\frac{2}{t} (1 - \cosh at)$
41	$\text{arc tan } \frac{\omega}{s}$	$\frac{1}{t} \sin \omega t$
42	$\frac{1}{s} \text{arc cot } s$	$Si(t)$ (cf. Sec. 17.2)

Chapter 5

Vector Analysis

Vectors are useful tools in engineering mathematics, because many physical quantities—for example, forces and velocities—may be represented by vectors, and in several respects the rules of vector calculation are as simple as the rules governing the system of real numbers.

In this chapter we shall consider the basic concepts and methods of vector analysis and their applications to various physical and geometrical problems.

It is true that any problem that can be solved by the use of vectors can also be treated by nonvectorial methods, but vector analysis is a shorthand which simplifies many calculations considerably. Furthermore, it is a way of visualizing physical and geometrical quantities and relations between them. For all those reasons extensive use is made of vector notation in modern technical literature.

The first part of the chapter (Secs. 5.1–5.8) is devoted to vector algebra. The foundations of vector calculus and its application to geometrical and mechanical problems are presented in Secs. 5.9–5.14. The remaining sections (5.15–5.19) include the notions of directional derivative, gradient, divergence, and curl and some of their most important applications.

Integral calculus in connection with vectors will be considered in the next chapter (Chap. 6).

Prerequisite for this chapter: determinants of the third order (Sec. 0.3). elementary differential calculus.

Sections which may be omitted in a shorter course: 5.8, 5.11–5.14, 5.17.

References: Appendix 1, Part C.

Answers to problems: Appendix 2.

5.1 Scalar and Vectors

In physics and geometry there are quantities each of which is completely specified when its magnitude—that is, its size or number of units according to some scale—is given. Examples are the mass of a body, the charge of an electron, the specific heat of water, the resistance of a resistor, the diameter of a circle, the area of a triangle, and the volume of a cube. Each of these

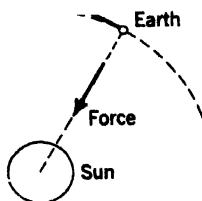
Velocity

Fig. 116. Force and velocity.

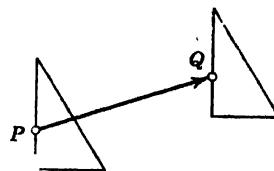


Fig. 117. Translation.

quantities is described by a single number (after a suitable choice of units of measure). Such a quantity is called a **scalar**.¹

However, there are other physical and geometrical quantities which cannot be described by a single number, because they require for their complete characterization the specification of the direction as well as the magnitude.

For example, forces in mechanics are quantities of this type. We know that we may represent a force graphically by an arrow, or *directed line segment*, which indicates the direction of the force and whose length is equal to the magnitude of the force according to some suitable scale. Figure 116 shows the force of attraction for the earth's motion around the sun. The instantaneous velocity of the earth may also be represented by an arrow of suitable length and direction, and this illustrates that a velocity is also a quantity which is characterized by a magnitude and a direction.

Figure 117 shows the translation (displacement without rotation) of a triangle in the plane. This motion can be characterized by its magnitude (distance travelled by each point of the triangle) and its direction. The translation may be indicated graphically by a directed line segment whose *initial point* is the original position P of a point of the triangle and whose *terminal point* is the new position Q of that point after the translation. If we do this for each point of the triangle, we obtain a family of directed line segments which have the *same length* and the *same direction* (that is, are parallel and are directed in the same sense). We may say that each of these directed line segments "carries" a point of the triangle from its original position to its new position.

This situation suggests the following definition.

Definition. *A directed line segment is called a vector. Its length is called the length or magnitude of the vector and its direction is called the direction of the vector. Two vectors are equal if, and only if, they have the same length and the same direction.*

Hence a vector may be arbitrarily translated (that is, displaced without rotation) or, what amounts to the same thing, its initial point may be chosen

¹ It is essential that the number which characterizes the scalar is independent of the choice of coordinates, but we shall discuss this point later (in Sec. 5.9).

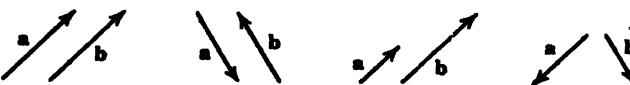


Fig. 118. Vectors.

arbitrarily. Clearly, if we choose a certain point as the initial point of a given vector, its terminal point is uniquely determined.

We shall denote vectors by lower-case boldfaced letters,² such as \mathbf{a} , \mathbf{b} , \mathbf{u} , \mathbf{v} , etc. The length of a vector v will be denoted by $|v|$.

If two vectors \mathbf{a} and \mathbf{b} are equal, we write

$$\mathbf{a} = \mathbf{b},$$

and if they are different, we may write

$$\mathbf{a} \neq \mathbf{b}.$$

Any vector may be represented graphically as an arrow of suitable length and direction, as is illustrated in Fig. 118.

A vector of length 1 is called a **unit vector**.

For the sake of completeness, we mention that in physics and geometry there are situations where we want to impose restrictions on the position of the initial point of a vector. For example, as is known from mechanics, a force acting on a rigid body may be equally well applied to any point of the body on its line of action. This suggests the concept of a **sliding vector**, that is, a vector whose initial point can be any point on a straight line which is parallel to the vector. A force acting on an elastic body is a vector whose initial point cannot be changed at all. In fact, if we choose another point of application of the force, its effect will in general be different. This suggests the notion of a **bound vector**, that is, a vector having a certain fixed initial point (*point of application*). Since these concepts occur in the literature, the student should know about them, but they will not be of particular importance in our further considerations.

5.2 Components of a Vector

We introduce a coordinate system in space whose axes are three mutually perpendicular straight lines. On all three axes we choose the same scale. Then the three **unit points** on the axes, whose coordinates are $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$, have the same distance from the **origin**, the point of intersection of the axes. The rectangular coordinate system thus obtained is called a **Cartesian coordinate system** in space (cf. Fig. 119).

We consider now a vector \mathbf{a} obtained by directing a line segment PQ such that P is the initial point and Q is the terminal point (Fig. 120).

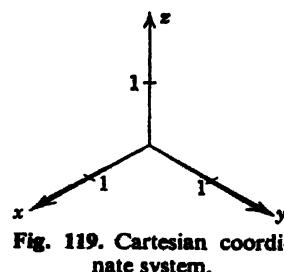


Fig. 119. Cartesian coordinate system.

² This is customary in printed work; in hand-written work one may characterize vectors by arrows, for example \vec{a} (in place of \mathbf{a}), \vec{b} , etc.

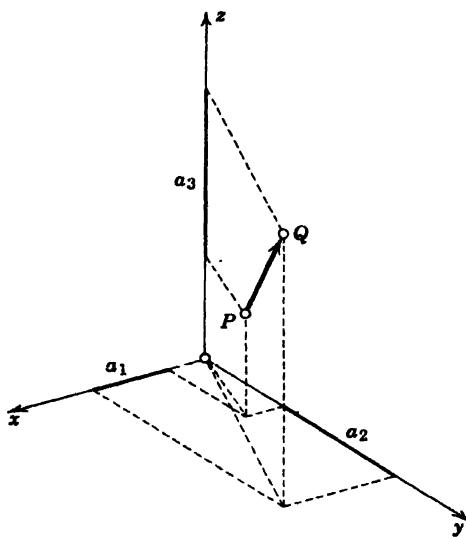


Fig. 120. Components of a vector.

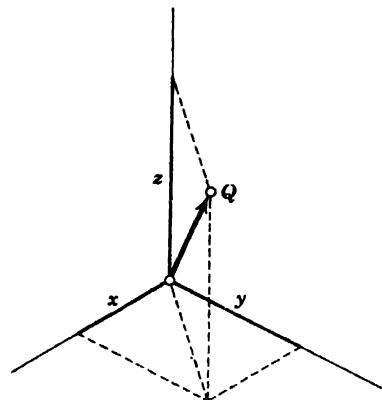


Fig. 121. Position vector.

Let (x_1, y_1, z_1) and (x_2, y_2, z_2) be the coordinates of P and Q , respectively. Then the numbers

$$(1) \quad a_1 = x_2 - x_1, \quad a_2 = y_2 - y_1, \quad a_3 = z_2 - z_1,$$

are called the **components** of the vector \mathbf{a} with respect to that Cartesian coordinate system.

By definition, the length $|\mathbf{a}|$ of the vector \mathbf{a} is the distance \overline{PQ} , and from (1) and the theorem of Pythagoras it follows that

$$(2) \quad |\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

Example 1. The vector \mathbf{a} with initial point $P: (3, 1, 4)$ and terminal point $Q: (1, -2, 4)$ has the components

$$a_1 = 1 - 3 = -2, \quad a_2 = -2 - 1 = -3, \quad a_3 = 4 - 4 = 0$$

and the length

$$|\mathbf{a}| = \sqrt{(-2)^2 + (-3)^2 + 0^2} = \sqrt{13}.$$

If we choose $(-1, 5, 8)$ as the initial point of \mathbf{a} , then the corresponding terminal point is $(-3, 2, 8)$.

From (1) we see immediately that the components a_1, a_2, a_3 of the vector \mathbf{a} are independent of the choice of the initial point of \mathbf{a} , because if we translate \mathbf{a} , then corresponding coordinates of P and Q are altered by the same amount. Hence, a fixed Cartesian coordinate system being given, each vector is uniquely determined by the ordered triple of its components with respect to that coordinate system.

If we introduce the **null vector or zero vector $\mathbf{0}$** , defined as the vector with components $0, 0, 0$, then any ordered triple of real numbers, including the

triple $0, 0, 0$, can be chosen as the components of a vector. The coordinate system being fixed, the correspondence between the ordered triples of real numbers and the vectors in space is one-to-one, that is, to each such triple there corresponds a vector in space that has those three numbers as its components with respect to that coordinate system, and conversely.

It follows that two vectors \mathbf{a} and \mathbf{b} are equal if, and only if, corresponding components of these vectors are equal. Consequently a vector equation

$$\mathbf{a} = \mathbf{b}$$

is equivalent to the three equations

$$a_1 = b_1, \quad a_2 = b_2, \quad a_3 = b_3$$

for the components a_1, a_2, a_3 and b_1, b_2, b_3 of the vectors with respect to a given Cartesian coordinate system.

A fixed Cartesian coordinate system being given, any point Q in space is uniquely determined by its coordinates x, y, z . Now x, y, z may be regarded as the components of a vector \mathbf{r} whose initial point is the origin of that coordinate system. Then Q is the terminal point of \mathbf{r} , and \mathbf{r} is called the **position vector** of Q with respect to that coordinate system (Fig. 121).

Problems

Find the components of the vector \mathbf{v} with given initial point $P: (x_1, y_1, z_1)$ and terminal point $Q: (x_2, y_2, z_2)$. Find $|\mathbf{v}|$. Graph \mathbf{v} .

- | | |
|------------------------------------|----------------------------------|
| 1. $P: (1, 2, 0), Q: (3, 2, 1)$ | 2. $P: (0, 1, -1), Q: (0, 4, 2)$ |
| 3. $P: (-1, 2, -2), Q: (0, 3, -1)$ | 4. $P: (1, 1, 1), Q: (3, 3, 3)$ |
| 5. $P: (-2, -3, -1), Q: (0, 0, 0)$ | 6. $P: (2, 4, 6), Q: (1, 2, 3)$ |
| 7. $P: (0, 2, 0), Q: (0, 1, 0)$ | 8. $P: (0, 0, 0), Q: (-1, 2, 0)$ |
| 9. $P: (3, 8, 2), Q: (3, 8, 2)$ | 10. $P: (3, 8, 4), Q: (3, 9, 4)$ |

In each case, the components v_1, v_2, v_3 of a vector \mathbf{v} and a particular initial point P are given. Find the corresponding terminal point.

- | | |
|---|--------------------------------|
| 11. $-1, 2, 5; P: (8, 2, 0)$ | 12. $3, 8, 0; P: (-4, 1, 2)$ |
| 13. $\frac{1}{2}, 2, -1; P: (-2, \frac{1}{2}, 0)$ | 14. $1, 0, 0; P: (0, 0, 0)$ |
| 15. $4, -1, 6; P: (-4, 1, -6)$ | 16. $0, 0, 0; P: (1, 4, 7)$ |
| 17. $3, 8, -9; P: (0, 0, 0)$ | 18. $1, 2, 3; P: (1, 2, 3)$ |
| 19. $-1, -2, 1; P: (-2, 0, -3)$ | 20. $8, -6, -1; P: (-8, 0, 4)$ |

5.3 Vector Addition. Multiplication by Scalars

Experiments show that the resultant of two forces can be determined by the familiar parallelogram law (Fig. 122). This suggests the following definition of vector addition.

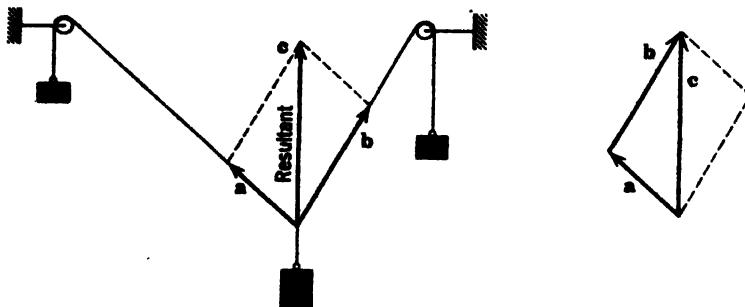


Fig. 122. Resultant of two forces (parallelogram law).

Given two vectors \mathbf{a} and \mathbf{b} , put the initial point of \mathbf{b} at the terminal point of \mathbf{a} ; then the sum of \mathbf{a} and \mathbf{b} is defined as the vector \mathbf{c} drawn from the initial point of \mathbf{a} to the terminal point of \mathbf{b} (Fig. 123), and we write

$$\mathbf{c} = \mathbf{a} + \mathbf{b}.$$

If in some fixed coordinate system \mathbf{a} has the components a_1, a_2, a_3 and \mathbf{b} has the components b_1, b_2, b_3 , then the components c_1, c_2, c_3 of the sum vector $\mathbf{c} = \mathbf{a} + \mathbf{b}$ are obtained by adding corresponding components of \mathbf{a} and \mathbf{b} ; thus

$$(1) \quad c_1 = a_1 + b_1, \quad c_2 = a_2 + b_2, \quad c_3 = a_3 + b_3.$$

From the definition or from (1) it follows that (Figs. 124, 125)

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}, \quad (\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}),$$

that is, vector addition is commutative and associative. Furthermore, for any vector \mathbf{a} ,

$$\mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a}.$$

Instead of $\mathbf{a} + \mathbf{a}$ we also write $2\mathbf{a}$ and so on. By $-\mathbf{a}$ we mean the vector whose direction is opposite to \mathbf{a} and whose length is $|\mathbf{a}|$. This suggests the following definition of multiplication of vectors by real numbers.

Let \mathbf{a} be any vector and q any real number. Then the vector $q\mathbf{a}$ is defined as follows.

The length of $q\mathbf{a}$ is $|q| |\mathbf{a}|$.

If $\mathbf{a} \neq \mathbf{0}$ and $q > 0$, then $q\mathbf{a}$ has the direction of \mathbf{a} .

If $\mathbf{a} \neq \mathbf{0}$ and $q < 0$, then $q\mathbf{a}$ has the direction opposite to \mathbf{a} .

If $\mathbf{a} = \mathbf{0}$ or $q = 0$ (or both), then $q\mathbf{a} = \mathbf{0}$.

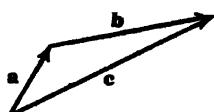


Fig. 123. Vector addition.

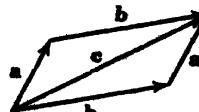


Fig. 124. Commutativity of vector addition.

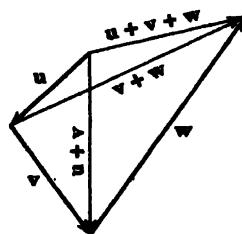


Fig. 125. Associativity of vector addition.

$$\begin{array}{c} \nearrow \\ 2\mathbf{a} \end{array} \quad \begin{array}{c} \nearrow \\ -\mathbf{a} \end{array} \quad \begin{array}{c} \nearrow \\ -\frac{1}{2}\mathbf{a} \end{array}$$

Fig. 126. Multiplication of vectors by numbers.

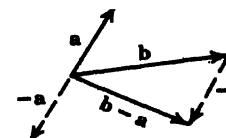


Fig. 127. Difference of vectors.

Simple examples are shown in Fig. 126.

Clearly, if \mathbf{a} has the components a_1, a_2, a_3 , then $q\mathbf{a}$ has the components qa_1, qa_2, qa_3 (with respect to the same coordinate system). Furthermore, from the definition it follows that

$$(c+k)\mathbf{a} = c\mathbf{a} + k\mathbf{a}, \quad c(k\mathbf{a}) = (ck)\mathbf{a} = cka, \quad q(\mathbf{a} + \mathbf{b}) = qa + qb,$$

$$1\mathbf{a} = \mathbf{a}, \quad (-1)\mathbf{a} = -\mathbf{a}, \quad \mathbf{a} + (-\mathbf{a}) = 0.$$

Instead of $\mathbf{b} + (-\mathbf{a})$ we simply write $\mathbf{b} - \mathbf{a}$ (Fig. 127).

A Cartesian coordinate system being given, we may now represent a vector \mathbf{a} with components a_1, a_2, a_3 as the sum of three vectors parallel to the coordinate axes. For this purpose we associate with that coordinate system three unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ which have the positive directions of the three coordinate axes. Then (Fig. 128)

$$(2) \quad \mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}.$$

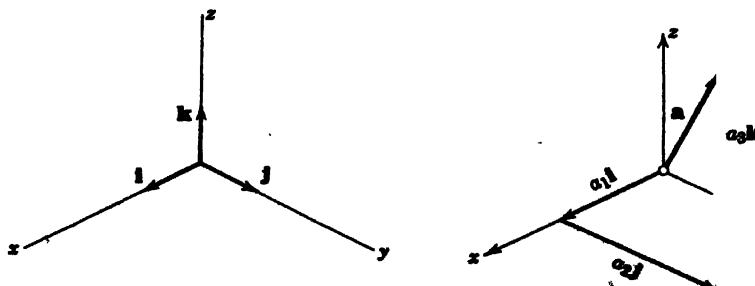
Figure 128 shows the vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ when the origin is chosen as their common initial point. These vectors are mutually perpendicular. Instead of *perpendicular* the term *orthogonal* is also used, and we say that $\mathbf{i}, \mathbf{j}, \mathbf{k}$ form a triple of orthogonal unit vectors. This triple is also known as the fundamental orthogonal triad associated with that coordinate system.

Example. With respect to a given coordinate system, let

$$\mathbf{a} = 4\mathbf{i} + \mathbf{k} \quad \text{and} \quad \mathbf{b} = 2\mathbf{i} - 4\mathbf{j} + 3\mathbf{k}.$$

Then

$$2\mathbf{b} = 4\mathbf{i} - 8\mathbf{j} + 6\mathbf{k}, \quad -\mathbf{b} = -2\mathbf{i} + 4\mathbf{j} - 3\mathbf{k}, \quad 1.5\mathbf{a} - \mathbf{b} = 4\mathbf{i} + 4\mathbf{j} - 1.5\mathbf{k}.$$

Fig. 128. The unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ and the representation (2).

Problems

Let $\mathbf{a} = \mathbf{i} + 2\mathbf{j} - 3\mathbf{k}$, $\mathbf{b} = 2\mathbf{i} + \mathbf{j} + 4\mathbf{k}$, $\mathbf{c} = -5\mathbf{j}$. Find

1. $-2\mathbf{a}$, $\frac{1}{2}\mathbf{a}$, $3\mathbf{a}$
2. $\mathbf{a} + \mathbf{b}$, $\mathbf{a} + \mathbf{b} + \mathbf{c}$
3. $\mathbf{a} - \mathbf{b}$, $\mathbf{b} - \mathbf{a}$, $\mathbf{b} - \mathbf{c}$
4. $|\mathbf{a} + \mathbf{b}|$, $|\mathbf{a}| + |\mathbf{b}|$
5. $|\mathbf{a} - \mathbf{c}|$, $|\mathbf{a}| - |\mathbf{c}|$
6. $\mathbf{a}/|\mathbf{a}|$, $\mathbf{b}/|\mathbf{b}|$, $\mathbf{c}/|\mathbf{c}|$
7. $3\mathbf{a} - 2\mathbf{b}$, $2\mathbf{b} - 3\mathbf{a}$
8. $|- \mathbf{a} + 2\mathbf{b} - 3\mathbf{c}|$
9. $|\mathbf{a} + \mathbf{b}| + |\mathbf{c}|$, $|\mathbf{a}| + |\mathbf{b} + \mathbf{c}|$

Find the resultant of the following forces.

10. $\mathbf{p} = \mathbf{i} + 3\mathbf{j} - \mathbf{k}$, $\mathbf{q} = 5\mathbf{i} - 2\mathbf{k}$, $\mathbf{u} = -\mathbf{j} + 3\mathbf{k}$
11. $\mathbf{p} = 3\mathbf{i} + 4\mathbf{j} + 2\mathbf{k}$, $\mathbf{q} = -2\mathbf{i} - 5\mathbf{j} - 3\mathbf{k}$, $\mathbf{u} = \mathbf{k}$
12. Determine a force \mathbf{p} such that the forces \mathbf{p} , $\mathbf{q} = \mathbf{i} + \mathbf{j} - 2\mathbf{k}$, and $\mathbf{u} = -\mathbf{i} + 2\mathbf{j}$ are in equilibrium.
13. Let \mathbf{a} , \mathbf{b} , \mathbf{c} be the edge vectors of a parallelepiped. Find the vectors corresponding to the four diagonals.

Using vectors, prove the following statements.

14. The diagonals of a rectangle have the same length.
15. The diagonals of a parallelogram bisect each other.
16. The four diagonals of a cube have the same length.
17. The line which joins one vertex of a parallelogram to the midpoint of an opposite side divides the diagonal in the ratio 1:2.
18. The sum of the vectors drawn from the center of a regular polygon to its vertices is the null vector.
19. The medians of a triangle meet at a point P which divides each median in the ratio 1:2.
20. The four diagonals of a parallelepiped meet and bisect each other.

5.4 Scalar Product

The scalar product (or dot product) of two vectors \mathbf{a} and \mathbf{b} is written $\mathbf{a} \cdot \mathbf{b}$ and is defined as

$$(1) \quad \begin{aligned} \mathbf{a} \cdot \mathbf{b} &= |\mathbf{a}| |\mathbf{b}| \cos \gamma && (\text{when } \mathbf{a} \neq 0, \mathbf{b} \neq 0), \\ \mathbf{a} \cdot \mathbf{b} &= 0 && (\text{when } \mathbf{a} = 0 \text{ or } \mathbf{b} = 0); \end{aligned}$$

here γ ($0 \leq \gamma \leq \pi$) is the angle between \mathbf{a} and \mathbf{b} (computed when the vectors have their initial point coinciding). (Cf. Fig. 129.)

The value of the scalar product is a scalar, a real number which is independent of the choice of Cartesian coordinate systems in space. Since the

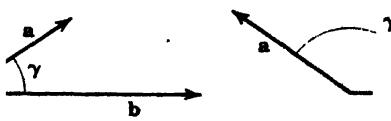


Fig. 129. Angle between vectors.

cosine in (1) may be negative or positive, the same is true for the scalar product. Since that cosine is zero if, and only if, $\gamma = \pi/2$, we obtain the following important result.

Theorem 1. *Two nonzero vectors are orthogonal if, and only if, their scalar product is zero.*

Furthermore, $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$ or

$$(2) \quad |\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}} \quad (\geq 0).$$

From this and (1) we obtain the useful formula

$$(3) \quad \cos \gamma = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| |\mathbf{b}|} = \frac{\mathbf{a} \cdot \mathbf{b}}{\sqrt{\mathbf{a} \cdot \mathbf{a}} \sqrt{\mathbf{b} \cdot \mathbf{b}}},$$

where γ is defined as before.

From the definition it follows readily that *scalar multiplication of vectors is commutative and is distributive with respect to vector addition*:

$$(4) \quad (a) \quad \mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}, \quad (b) \quad \mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}.$$

Let \mathbf{a} and \mathbf{b} ($\neq 0$) be given vectors and let γ denote the angle between them. Then the real number

$$p = |\mathbf{a}| \cos \gamma$$

is called the **component** of \mathbf{a} in the direction of \mathbf{b} . If $\mathbf{a} = 0$, then γ is undefined, and we set $p = 0$.

It follows that $|p|$ is the length of the orthogonal projection of \mathbf{a} on a straight line l in the direction of \mathbf{b} . p may be positive, zero, or negative (Fig. 130).

From this definition we see that in particular the components of a vector \mathbf{a} in the directions of the unit vectors \mathbf{i} , \mathbf{j} , \mathbf{k} of the fundamental triad associated with a Cartesian coordinate system are the components a_1 , a_2 , a_3 of \mathbf{a} as defined in Sec. 5.2. This shows that our present use of the term "component" is merely a slight generalization of the previous one.

From (3) we obtain

$$(5) \quad p = |\mathbf{a}| \cos \gamma = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{b}|} \quad (\mathbf{b} \neq 0)$$

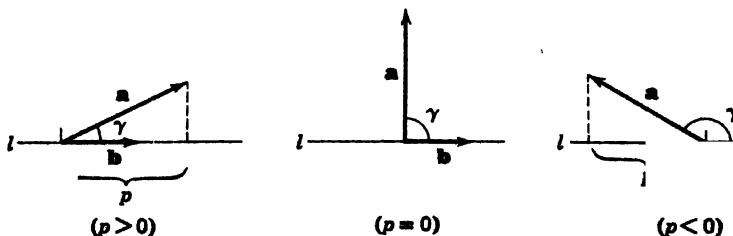


Fig. 130. Component of a vector \mathbf{a} in the direction of a vector \mathbf{b} .

and if in particular \mathbf{b} is a unit vector, then we simply have

$$(6) \quad p = \mathbf{a} \cdot \mathbf{b}.$$

Given a fixed Cartesian coordinate system, we may easily express the scalar product $\mathbf{a} \cdot \mathbf{b}$ of two vectors

$$(7) \quad \mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} \quad \text{and} \quad \mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$$

in terms of the components of these vectors. Since \mathbf{i} , \mathbf{j} , and \mathbf{k} are unit vectors, we have

$$\mathbf{i} \cdot \mathbf{i} = 1, \quad \mathbf{j} \cdot \mathbf{j} = 1, \quad \mathbf{k} \cdot \mathbf{k} = 1,$$

and since they are orthogonal, it follows from Theorem 1 that

$$\mathbf{i} \cdot \mathbf{j} = 0, \quad \mathbf{j} \cdot \mathbf{k} = 0, \quad \mathbf{k} \cdot \mathbf{i} = 0.$$

Hence, if we substitute (7) into $\mathbf{a} \cdot \mathbf{b}$ and use the distributive law (4b), we first have a sum of nine scalar products,

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1\mathbf{i} \cdot \mathbf{i} + a_1b_2\mathbf{i} \cdot \mathbf{j} + \cdots + a_3b_3\mathbf{k} \cdot \mathbf{k}.$$

Since six of these products are zero, we obtain the simple basic formula

$$(8) \quad \mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3.$$

The following examples may illustrate the usefulness of scalar products. Various other applications will be considered later.

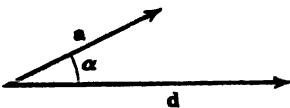


Fig. 131. Work done by a force.

Example 1 (Work done by a force). Consider a particle on which a constant force \mathbf{a} acts. Let the particle be given a displacement \mathbf{d} . Then the work W done by \mathbf{a} in the displacement is defined as the product of $|\mathbf{d}|$ and the component of \mathbf{a} in the direction of \mathbf{d} , that is,

$$(9) \quad W = |\mathbf{a}| |\mathbf{d}| \cos \alpha = \mathbf{a} \cdot \mathbf{d}$$

where α is the angle between \mathbf{d} and \mathbf{a} (Fig. 131).

Example 2 (Orthogonal straight lines in a plane). Find a representation of the straight line L_1 through the point $P: (1, 3)$ in the xy -plane and perpendicular to the line L_2 represented by $x - 2y + 2 = 0$.

Any straight line L_1 in the xy -plane can be represented in the form $a_1x + a_2y = c$. If $c = 0$, then L_1 passes through the origin. If $c \neq 0$, then $a_1x + a_2y = c$ represents a line L_1^* through the origin and parallel to L_1 . The position vector of a point on L_1^* is $\mathbf{r} = xi + yj$. Introducing the vector $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j}$, we see from (8) that the representation of L_1^* may be written

$$\mathbf{a} \cdot \mathbf{r} = 0.$$

Certainly $\mathbf{a} \neq 0$ and, by Theorem 1, the vector \mathbf{a} is perpendicular to \mathbf{r} and, therefore, perpendicular to the line L_1^* . It is called a normal vector to L_1^* . Since L_1 and L_1^* are parallel, \mathbf{a} is also a normal vector to L_1 . Hence two lines L_1 and L_2 : $b_1x + b_2y = d$ are perpendicular or orthogonal if, and only if, their normal vectors \mathbf{a} and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j}$ are orthogonal, that is, $\mathbf{a} \cdot \mathbf{b} = 0$. Note that this implies that the slopes of the lines are negative reciprocals.

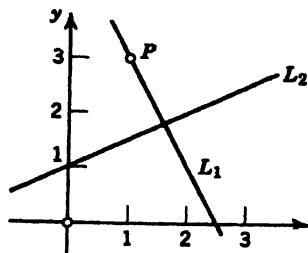


Fig. 132. Example 2.

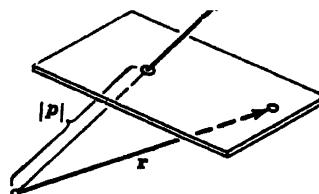


Fig. 133. Normal vector to a plane.

In our case, $\mathbf{b} = \mathbf{i} - 2\mathbf{j}$, and a vector perpendicular to \mathbf{b} is $\mathbf{a} = 2\mathbf{i} + \mathbf{j}$. Hence the representation of L_1 must be of the form $2x + y = c$, and by substituting the coordinates of P we obtain $c = 5$. This yields the solution

$$y = -2x + 5.$$

Example 3 (Normal vector to a plane). Find a unit vector perpendicular to the plane $4x + 2y + 4z = -7$.

Any plane in space can be represented in the form

$$(10) \quad a_1x + a_2y + a_3z = c.$$

The position vector of a point in the plane is $\mathbf{r} = xi + yj + zk$. Introducing the vector $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and using (8), we may write (10) in the form

$$(11) \quad \mathbf{a} \cdot \mathbf{r} = c.$$

Certainly $\mathbf{a} \neq 0$, and the unit vector in the direction of \mathbf{a} is

$$\mathbf{n} = \frac{\mathbf{a}}{|\mathbf{a}|}.$$

Dividing by $|\mathbf{a}|$, we obtain from (11)

$$(12) \quad \mathbf{n} \cdot \mathbf{r} = p \quad \text{where} \quad p = \frac{c}{|\mathbf{a}|}.$$

From (6) we see that p is the projection of \mathbf{r} in the direction of \mathbf{n} , and this projection has the same constant value $c/|\mathbf{a}|$ for the position vector \mathbf{r} of any point in the plane. Clearly this holds if, and only if, \mathbf{n} is perpendicular to the plane. \mathbf{n} is called a unit normal vector to the plane (the other being $-\mathbf{n}$). Furthermore from this and the definition of projection it follows that $|p|$ is the distance of the plane from the origin. Representation (12) is called Hesse's³ normal form of a plane.

In our case,

$$\mathbf{a} = 4\mathbf{i} + 2\mathbf{j} + 4\mathbf{k}, \quad c = -7, \quad |\mathbf{a}| = 6,$$

$$\mathbf{n} = \frac{1}{6}\mathbf{a} = \frac{1}{3}\mathbf{i} + \frac{1}{3}\mathbf{j} + \frac{2}{3}\mathbf{k},$$

and the plane has the distance $7/6$ from the origin.

³ LUDWIG OTTO HESSE (1811–1874), German mathematician, who contributed to the theory of curves and surfaces.

Problems

Let $\mathbf{a} = \mathbf{i} + 3\mathbf{j} + 2\mathbf{k}$, $\mathbf{b} = -\mathbf{j} + 4\mathbf{k}$, and $\mathbf{c} = 3\mathbf{i} - 4\mathbf{j} - \mathbf{k}$. Find

- | | | |
|--|--|---|
| 1. $\mathbf{a} \cdot \mathbf{b}$, $\mathbf{b} \cdot \mathbf{a}$, $\mathbf{b} \cdot \mathbf{c}$ | 2. $ \mathbf{a} $, $ \mathbf{b} $, $ \mathbf{c} $ | 3. $(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c}$, $\mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}$ |
| 4. $\mathbf{a} \cdot (\mathbf{b} - \mathbf{c})$, $\mathbf{a} \cdot (\mathbf{c} - \mathbf{b})$ | 5. $ \mathbf{a} + \mathbf{b} $, $ \mathbf{b} + \mathbf{c} $ | 6. $2\mathbf{a} \cdot 3\mathbf{b}$, $6\mathbf{a} \cdot \mathbf{b}$ |
| 7. $ \mathbf{a} + \mathbf{c} $, $ \mathbf{a} + \mathbf{c} $ | 8. $2\mathbf{a} \cdot (\mathbf{a} - 2\mathbf{b})$ | 9. $-\mathbf{a} \cdot (2\mathbf{c} - \mathbf{b})$ |

Find the work done by a force \mathbf{p} acting on a particle, if the particle is displaced from a point A to a point B along a straight segment AB , where

10. $A: (0, 0, 0)$, $B: (0, 3, 0)$, $\mathbf{p} = \mathbf{i} + \mathbf{j} + \mathbf{k}$
11. $A: (0, -1, 2)$, $B: (0, 0, 0)$, $\mathbf{p} = \mathbf{k}$
12. $A: (8, -2, -3)$, $B: (-2, 0, 6)$, $\mathbf{p} = 3\mathbf{i} - 2\mathbf{j} + 4\mathbf{k}$
13. $A: (2, 2, 2)$, $B: (4, 4, 4)$, $\mathbf{p} = -4\mathbf{j} + 2\mathbf{k}$
14. $A: (4, 1, 3)$, $B: (5, 2, 5)$, $\mathbf{p} = \mathbf{i} + \mathbf{j} - \mathbf{k}$
15. Why is the work in Prob. 14 equal to zero?
16. For what values of a_1 are $\mathbf{a} = a_1\mathbf{i} + 2\mathbf{j}$ and $\mathbf{b} = 3\mathbf{i} + 4\mathbf{j} - \mathbf{k}$ orthogonal vectors?
17. Find a unit vector $a_1\mathbf{i} + a_2\mathbf{j}$ perpendicular to the vector $2\mathbf{i} - 5\mathbf{j}$.
18. Show that the vectors $\mathbf{a} = \mathbf{i} + \mathbf{j} + \mathbf{k}$, $\mathbf{b} = \mathbf{i} + 2\mathbf{j} - 3\mathbf{k}$, and $\mathbf{c} = -5\mathbf{i} + 4\mathbf{j} + \mathbf{k}$ are orthogonal.
19. Are the four diagonals of a cube orthogonal?
20. Show that the planes $x + y + z = 1$ and $3x - 4y + z = 2$ are orthogonal.

Let $\mathbf{a} = 3\mathbf{i} + \mathbf{j} + 4\mathbf{k}$, $\mathbf{b} = \mathbf{i} - \mathbf{j}$, and $\mathbf{c} = \mathbf{i} + 5\mathbf{k}$. Find the cosine of the angle between the following vectors.

21. \mathbf{a}, \mathbf{c}
22. $\mathbf{a}, -\mathbf{c}$
23. $\mathbf{a}, \mathbf{b} + \mathbf{c}$
24. $\mathbf{a} + \mathbf{b}, \mathbf{a} - \mathbf{b}$
25. Show that if $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are orthogonal unit vectors, then $\mathbf{a} + \mathbf{b} + \mathbf{c}$ makes the same angle with each of the vectors.

Find the angles of the triangle whose vertices are

26. $A: (0, 0, 0)$, $B: (4, -1, 3)$, $C: (1, 2, 3)$
27. $A: (1, 1, 0)$, $B: (3, 1, 0)$, $C: (1, 3, 0)$
28. Find the angle between the planes $x + 2y + z = 2$ and $2x - y + 3z = -4$.
29. In Fig. 134, $\mathbf{c} = \mathbf{a} - \mathbf{b}$. Hence, $|\mathbf{c}|^2 = (\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b})$. Deduce the law of cosines from this relation.

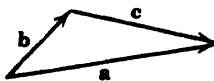


Fig. 134. Problem 29.

30. Let $\mathbf{a} = \cos \alpha \mathbf{i} + \sin \alpha \mathbf{j}$, and $\mathbf{b} = \cos \beta \mathbf{i} + \sin \beta \mathbf{j}$ where $0 \leq \alpha \leq \beta \leq 2\pi$. Show that \mathbf{a} and \mathbf{b} are unit vectors. Use (8) to obtain the trigonometric identity for $\cos(\beta - \alpha)$.

5.5 Vector Product

Various applications suggest the introduction of another kind of vector multiplication in which the product of two vectors is again a vector. This so-called **vector product** or **cross product** of two vectors \mathbf{a} and \mathbf{b} is written

$$\mathbf{a} \times \mathbf{b}$$

and is a vector \mathbf{v} which is defined as follows.

If \mathbf{a} and \mathbf{b} have the same or opposite direction or one of these vectors is the zero vector, then $\mathbf{v} = \mathbf{0}$.

In any other case, \mathbf{v} is the vector whose length is equal to the area of the parallelogram with \mathbf{a} and \mathbf{b} as adjacent sides and whose direction is perpendicular to both \mathbf{a} and \mathbf{b} and is such that $\mathbf{a}, \mathbf{b}, \mathbf{v}$, in this order, form a right-handed triple or right-handed triad, as shown in Fig. 135.

The term **right-handed** comes from the fact that the vectors $\mathbf{a}, \mathbf{b}, \mathbf{v}$, in this order, assume the same sort of orientation as the thumb, index finger, and middle finger of the right hand when these are held as shown in Fig. 136. We may also say that if \mathbf{a} is rotated into the direction of \mathbf{b} through the angle α ($<\pi$), then \mathbf{v} advances in the same direction as a right-handed screw would if turned in the same way (cf. Fig. 137).

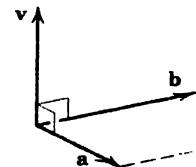


Fig. 135. Vector product.



Fig. 136. Right-handed triple of vectors $\mathbf{a}, \mathbf{b}, \mathbf{v}$.

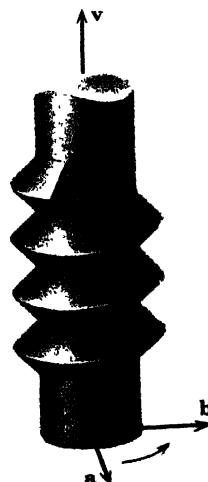


Fig. 137. Right-handed screw.

The parallelogram with \mathbf{a} and \mathbf{b} as adjacent sides has the area $|\mathbf{a}| |\mathbf{b}| \sin \gamma$ (cf. Fig. 135). We thus obtain

$$(1) \quad |\mathbf{v}| = |\mathbf{a}| |\mathbf{b}| \sin \gamma.$$

Let $\mathbf{a} \times \mathbf{b} = \mathbf{v}$ and let $\mathbf{b} \times \mathbf{a} = \mathbf{w}$. Then, by definition, $|\mathbf{v}| = |\mathbf{w}|$, and in order that $\mathbf{b}, \mathbf{a}, \mathbf{w}$ form a right-handed triple we must have $\mathbf{w} = -\mathbf{v}$ (Fig. 138). Hence,

$$(2) \quad \mathbf{b} \times \mathbf{a} = -(\mathbf{a} \times \mathbf{b});$$

that is, *cross-multiplication of vectors is not commutative but anticommutative*. The order of the factors in a vector product is, therefore, of great importance and must be carefully observed.

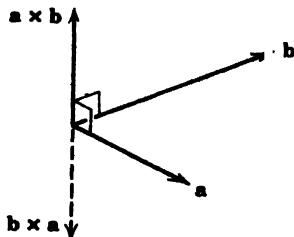


Fig. 138. Anticommutativity of cross multiplication.

From the definition it follows that for any constant k ,

$$(3) \quad (k\mathbf{a}) \times \mathbf{b} = k(\mathbf{a} \times \mathbf{b}) = \mathbf{a} \times (k\mathbf{b}).$$

Furthermore cross-multiplication is distributive with respect to vector addition, that is,

$$(4) \quad \begin{aligned} \mathbf{a} \times (\mathbf{b} + \mathbf{c}) &= (\mathbf{a} \times \mathbf{b}) + (\mathbf{a} \times \mathbf{c}), \\ (\mathbf{a} + \mathbf{b}) \times \mathbf{c} &= (\mathbf{a} \times \mathbf{c}) + (\mathbf{b} \times \mathbf{c}). \end{aligned}$$

The proof will be given in the next section. More complicated cross products will be considered later, but we want to mention now that

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c},$$

ordinarily.

From (1) in this section and (1) in the last section it follows that

$$|\mathbf{v}|^2 = |\mathbf{a}|^2 |\mathbf{b}|^2 \sin^2 \gamma = |\mathbf{a}|^2 |\mathbf{b}|^2 (1 - \cos^2 \gamma) = (\mathbf{a} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{b}) - (\mathbf{a} \cdot \mathbf{b})^2.$$

Hence we obtain the formula

$$(5) \quad |\mathbf{a} \times \mathbf{b}| = \sqrt{(\mathbf{a} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{b}) - (\mathbf{a} \cdot \mathbf{b})^2},$$

which is convenient for computing the length of a vector product.

5.6 Vector Products in Terms of Components

We shall now represent a vector product in terms of the components of its factors with respect to a Cartesian coordinate system. In this connection it is important to note that there are two types of such systems, depending on the orientation of the axes, namely, right-handed and left-handed. The definitions are as follows.

A Cartesian coordinate system is called **right-handed**, if the corresponding unit vectors \mathbf{i} , \mathbf{j} , \mathbf{k} in the positive directions of the axes form a right-handed triple (Fig. 139a); it is called **left-handed**, if these vectors form a *left-handed triple*, that is, assume the same sort of orientation as the thumb, index finger, and middle finger of the left hand (Fig. 139b).

In applications we preferably use right-handed systems.

With respect to a given right-handed or left-handed Cartesian coordinate system, let a_1, a_2, a_3 and b_1, b_2, b_3 be the components of two vectors \mathbf{a} and \mathbf{b} , respectively. We want to express the components v_1, v_2, v_3 of the product vector

$$\mathbf{v} = \mathbf{a} \times \mathbf{b}$$

in terms of the components of \mathbf{a} and \mathbf{b} . We need only consider the case when $\mathbf{v} \neq 0$. Since \mathbf{v} is perpendicular to both \mathbf{a} and \mathbf{b} , it follows from Theorem 1, Sec. 5.4, that $\mathbf{a} \cdot \mathbf{v} = 0$ and $\mathbf{b} \cdot \mathbf{v} = 0$ or, by (8) in Sec. 5.4,

$$(1) \quad \begin{aligned} a_1 v_1 + a_2 v_2 + a_3 v_3 &= 0, \\ b_1 v_1 + b_2 v_2 + b_3 v_3 &= 0. \end{aligned}$$

Multiplying the first equation by b_3 , the last by a_3 , and subtracting, we obtain

$$(a_3 b_1 - a_1 b_3) v_1 = (a_2 b_3 - a_3 b_2) v_2.$$

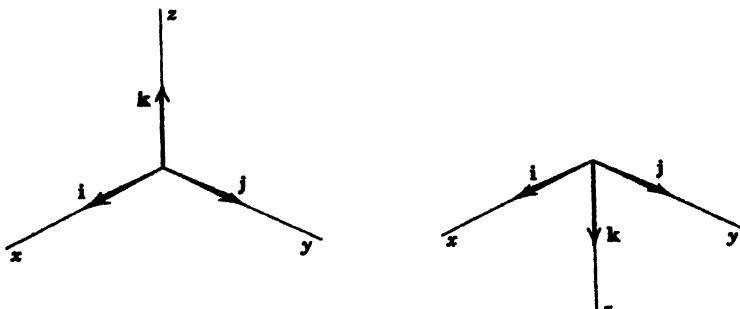
Multiplying the first equation of (1) by b_1 , the last by a_1 , and subtracting, we obtain

$$(a_1 b_2 - a_2 b_1) v_2 = (a_3 b_1 - a_1 b_3) v_3.$$

We readily see that these two equations are satisfied by

$$(2) \quad v_1 = c(a_2 b_3 - a_3 b_2), \quad v_2 = c(a_3 b_1 - a_1 b_3), \quad v_3 = c(a_1 b_2 - a_2 b_1),$$

where c is a constant. The reader may verify by inserting that (2) also satisfies (1). Now each of the equations in (1) represents a plane through the origin



(a) Right-handed.

(b) Left-handed.

Fig. 139. The two types of Cartesian coordinate systems.

in $v_1 v_2 v_3$ -space. The vectors \mathbf{a} and \mathbf{b} are normal vectors of these planes (cf. Ex. 3 in Sec. 5.4). Since $\mathbf{v} \neq 0$, these vectors are not parallel and the two planes do not coincide. Hence their intersection is a straight line L through the origin. Since (2) is a solution of (1) and, for varying c , represents a straight line, we conclude that (2) represents L , and every solution of (1) must be of the form (2). In particular, the components of \mathbf{v} must be of this form, where c is to be determined. From (2) we obtain

$$|\mathbf{v}|^2 = v_1^2 + v_2^2 + v_3^2 = c^2[(a_2 b_3 - a_3 b_2)^2 + (a_3 b_1 - a_1 b_3)^2 + (a_1 b_2 - a_2 b_1)^2],$$

and we may readily verify that this may be written

$$|\mathbf{v}|^2 = c^2[(a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1 b_1 + a_2 b_2 + a_3 b_3)^2].$$

Using (8), Sec. 5.4, we thus have

$$|\mathbf{v}|^2 = c^2[(\mathbf{a} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{b}) - (\mathbf{a} \cdot \mathbf{b})^2].$$

By comparing this with (5) in the last section we conclude that $c = \pm 1$.

From now on it will matter whether the coordinate system is right-handed or left-handed. We first consider the case of a right-handed system, and we want to show that then $c = +1$. This can be done as follows.

If we change the lengths and directions of \mathbf{a} and \mathbf{b} continuously and such that at the end $\mathbf{a} = \mathbf{i}$ and $\mathbf{b} = \mathbf{j}$ (Fig. 139a), then \mathbf{v} will change its length and direction continuously, and at the end, $\mathbf{v} = \mathbf{i} \times \mathbf{j} = \mathbf{k}$. Obviously we may effect the change so that both \mathbf{a} and \mathbf{b} remain different from the zero vector and are not parallel at any instant. Then \mathbf{v} is never equal to the zero vector, and since the change is continuous and c can only assume the values $+1$ or -1 , it follows that at the end c must have the same value as before. Now at the end $\mathbf{a} = \mathbf{i}$, $\mathbf{b} = \mathbf{j}$, $\mathbf{v} = \mathbf{k}$ and, therefore, $a_1 = 1$, $b_2 = 1$, $v_3 = 1$, and the other components in (2) are zero. Hence, from (2), $v_3 = c = +1$. Noting that the expressions in parentheses in (2) may be written in the form of second-order determinants, we may sum up our result as follows.

With respect to a right-handed Cartesian coordinate system,

$$(3) \quad \mathbf{a} \times \mathbf{b} = \mathbf{i} \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} + \mathbf{j} \begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix} + \mathbf{k} \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix},$$

where a_1, a_2, a_3 and b_1, b_2, b_3 are the components of \mathbf{a} and \mathbf{b} , respectively.

For memorizing, it is useful to note that (3) can be interpreted as the expansion of the determinant

$$(4) \quad \mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

by the first row, but we should keep in mind that this is not an ordinary determinant because the elements of the first row are vectors.

In a *left-handed* Cartesian coordinate system, $\mathbf{i} \times \mathbf{j} = -\mathbf{k}$ (Fig. 139b), and the above reasoning leads to $c = -1$ in (2). Hence with respect to a left-handed Cartesian coordinate system,

$$(5) \quad \mathbf{a} \times \mathbf{b} = - \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$

Example 1. With respect to a right-handed Cartesian coordinate system, let $\mathbf{a} = 4\mathbf{i} - \mathbf{k}$ and $\mathbf{b} = -2\mathbf{i} + \mathbf{j} + 3\mathbf{k}$. Then

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 4 & 0 & -1 \\ -2 & 1 & 3 \end{vmatrix} = \mathbf{i} - 10\mathbf{j} + 4\mathbf{k}.$$

We may now prove the distributive law (4) in the previous section. From (3) it follows that the first component of $\mathbf{a} \times (\mathbf{b} + \mathbf{c})$ is

$$\begin{aligned} \begin{vmatrix} a_2 & a_3 \\ b_2 + c_2 & b_3 + c_3 \end{vmatrix} &= a_2(b_3 + c_3) - a_3(b_2 + c_2) \\ &= (a_2b_3 - a_3b_2) + (a_2c_3 - a_3c_2) \\ &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} + \begin{vmatrix} a_2 & a_3 \\ c_2 & c_3 \end{vmatrix}. \end{aligned}$$

The expression on the right is the first component of $\mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$. For the other two components of that vector the consideration is similar. This proves the first of the relations (4) in the previous section, and the last one can be proved by the same argument.

The definition of cross-multiplication is suggested by many applications, two of which may be illustrated by the following examples. Further physical and geometrical applications will be considered later.

Example 2 (Moment of a force). In mechanics the moment m of a force \mathbf{p} about a point Q is defined as the product $m = |\mathbf{p}| d$ where d is the (perpendicular) distance between Q and the line of action L of \mathbf{p} (Fig. 140). If \mathbf{r} is the vector from Q to any point A on L , then $d = |\mathbf{r}| \sin \gamma$ (Fig. 140) and

$$m = |\mathbf{r}| |\mathbf{p}| \sin \gamma.$$

Since γ is the angle between \mathbf{r} and \mathbf{p} ,

$$m = |\mathbf{r} \times \mathbf{p}|,$$

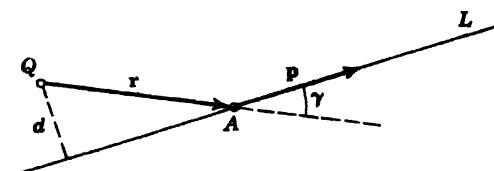


Fig. 140. Moment of a force.

as follows from (1) in the last section. The vector

$$(6) \quad m = r \times p$$

is called the moment vector or vector moment of p about Q . Its magnitude is m , and its direction is that of the axis of the rotation about Q which p has the tendency to produce.

Example 3 (Velocity of a rotating body). A rotation of a rigid body B in space can be simply and uniquely described by a vector w as follows. The direction of w is that of the axis of rotation and such that the rotation appears clockwise, if one looks from the initial point of w to its terminal point. The length of w is equal to the angular speed $\omega (>0)$ of the rotation, that is, the linear (or tangential) speed of a point of B divided by its distance from the axis of rotation.

Let P be any point of B and d its distance from the axis. Then P has the speed ωd . Let r be the position vector of P referred to a coordinate system with origin O on the axis of rotation. Then $d = |r| \sin \gamma$ where γ is the angle between w and r . Therefore,

$$\omega d = |w| |r| \sin \gamma = |w \times r|.$$

From this and the definition of vector product we see that the velocity vector v of P can be represented in the form (Fig. 141)

$$(7) \quad v = w \times r.$$

This simple formula is useful for determining v at any point of B .

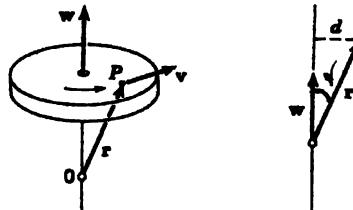


Fig. 141. Rotation of a rigid body.

Problems

With respect to a right-handed Cartesian coordinate system, let $a = i + 2j - 3k$, $b = i + 2j$, and $c = -i + j$. Find

- | | | |
|--------------------------------|----------------------------------|-------------------------------|
| 1. $a \times b$, $b \times a$ | 2. $b \times c$, $ b \times c $ | 3. $a \times c$, $a \cdot c$ |
| 4. $(a + 2b) \times c$ | 5. $(c - a) \times 2b$ | 6. $a \times (b + c)$ |
| 7. $a \times (2b + 3c)$ | 8. $(a + b) \times c$ | 9. $a \times (b - c)$ |
| 10. $\frac{1}{2}a \times 3b$ | 11. $a \times a$ | 12. $(a - 3b) \times b$ |
| 13. $(a - b) \times (b - c)$ | 14. $(a \times b) \times c$ | 15. $a \times (b \times c)$ |

Using (5), Sec. 5.5, find $|a \times b|$, where

- | | |
|-------------------------------------|--------------------------------------|
| 16. $a = i + j - k$, $b = 3k$ | 17. $a = 2i + 4k$, $b = 5j - k$ |
| 18. $a = 3i - 4j$, $b = i - j - k$ | 19. $a = 5i + 2j + k$, $b = 2j - k$ |

Find the area of the parallelogram of which the given vectors are adjacent sides.

20. $\mathbf{i} + \mathbf{j}, \mathbf{i} - \mathbf{j}$

21. $3\mathbf{j} - \mathbf{k}, \mathbf{i} + 2\mathbf{j}$

22. $\mathbf{i} + 2\mathbf{j} - \mathbf{k}, \mathbf{j} + \mathbf{k}$

23. $4\mathbf{i} - \mathbf{j} - \mathbf{k}, \mathbf{i} + 2\mathbf{j}$

24. $\mathbf{i}, \mathbf{i} + \mathbf{j} + \mathbf{k}$

25. $\mathbf{i} - \mathbf{j} + \mathbf{k}, \mathbf{i} + \mathbf{j} - \mathbf{k}$

Find the area of the parallelogram that has the following vertices in the xy -plane.

26. $(0, 0), (1, 0), (1, 1), (2, 1)$

27. $(1, 1), (4, 1), (2, 2), (5, 2)$

28. $(0, 0), (2, 2), (-1, 1), (1, 3)$

29. $(1, 2), (0, 0), (2, 6), (1, 4)$

30. $(-4, 2), (-6, 5), (-3, 6), (-5, 9)$

31. $(8, -3), (10, 1), (9, 0), (11, 4)$

Find the area of the triangle that has the following vertices.

32. $(0, 0, 0), (0, 1, 0), (1, 1, 0)$

33. $(3, 2, 0), (1, -1, 0), (2, 3, 0)$

34. $(0, 4, 0), (1, 1, 1), (-2, 1, -3)$

35. $(4, -2, 6), (6, -1, 7), (5, 0, 5)$

Are the following vectors parallel? Perpendicular?

36. $-2\mathbf{i} + 4\mathbf{j}, 3\mathbf{k}$

37. $3\mathbf{i} + 2\mathbf{j} + \mathbf{k}, 6\mathbf{i} + 4\mathbf{j} + 2\mathbf{k}$

38. $\mathbf{i} + \mathbf{j}, \mathbf{i} - \mathbf{j}$

39. $\mathbf{i} - 2\mathbf{j} + 3\mathbf{k}, 3\mathbf{i} + 5\mathbf{j}$

40. $-\mathbf{i} + \mathbf{j} - 3\mathbf{k}, 4\mathbf{i} - 4\mathbf{j} + 12\mathbf{k}$

41. $\mathbf{i} + \mathbf{j} - 2\mathbf{k}, -3\mathbf{i} - 3\mathbf{j} + 6\mathbf{k}$

Find two unit vectors perpendicular to both vectors.

42. \mathbf{j}, \mathbf{k}

43. \mathbf{i}, \mathbf{j}

44. $\mathbf{i} + \mathbf{j} + \mathbf{k}, \mathbf{i}$

45. $\mathbf{j} + \mathbf{k}, \mathbf{j} - \mathbf{k}$

46. $\mathbf{i} - 2\mathbf{j} + \mathbf{k}, \mathbf{k}$

47. $2\mathbf{i} - 2\mathbf{j}, 2\mathbf{j} + \mathbf{k}$

Find a unit normal vector to the plane through the following points.

48. $(0, 0, 0), (0, 1, 0), (0, 0, 1)$

49. $(-1, 2, 3), (1, 1, 1), (2, -1, 3)$

50. $(0, -1, 0), (2, 2, 1), (1, 0, 1)$

51. $(3, 2, 8), (1, 3, 3), (0, 2, 1)$

52. Find a vector \mathbf{v} parallel to the intersection of the planes $2x + 3y + 4z = 0$ and $x - y + z = 2$.

53. Find a vector \mathbf{v} parallel to the plane $x + y + z = 1$ and perpendicular to the line $y = x, z = 0$.

54. A sphere of radius 4 cm with center at the origin rotates about the z -axis with angular velocity $\omega = 2$ radians per second such that the rotation appears counter-clockwise if we look in the positive z -direction. Find the corresponding rotation vector \mathbf{w} . Find the velocity vector \mathbf{v} of a point P on the "equator" of the sphere at the instant when P has the position $\sqrt{8}\mathbf{i} - \sqrt{8}\mathbf{j}$ in space. The coordinates are right-handed.

A force \mathbf{p} acts on a line through a point A . Find the moment \mathbf{m} of \mathbf{p} about a point Q , where

55. $\mathbf{p} = \mathbf{i} + \mathbf{j}, A: (0, 0, 0), Q: (0, 1, 0)$

56. $\mathbf{p} = \mathbf{i}, A: (1, 1, 0), Q: (-5, 1, 0)$

57. $\mathbf{p} = 2\mathbf{i} + \mathbf{j}, A: (1, 1, 1), Q: (2, 2, 2)$

58. $\mathbf{p} = 3\mathbf{i} - \mathbf{j} + 2\mathbf{k}, A: (0, -1, 4), Q: (3, 0, 2)$

59. If $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ and $\mathbf{b} = \mathbf{a} \times \mathbf{c}$, show that $\mathbf{b} = \mathbf{0}$ and $\mathbf{c} = \mathbf{0}$.

60. If $\mathbf{v} = \mathbf{a} \times \mathbf{b}$, show that $\mathbf{a} \cdot \mathbf{v} = 0$ and $\mathbf{b} \cdot \mathbf{v} = 0$.

5.7 Scalar Triple Product. Linear Dependence of Vectors

Repeated products of vectors that have three or more factors occur frequently in applications. The most important of these products is the **scalar triple product** or *mixed triple product* $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ of three vectors. With respect to any right-handed Cartesian coordinate system, let

$$\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}, \quad \mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}, \quad \mathbf{c} = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}.$$

Then from (4), Sec. 5.6, it follows that

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) : (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \cdot \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

From this and (8), Sec. 5.4, we see that the scalar triple product takes the form

$$(1) \quad \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) : \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

The scalar triple product $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ will be denoted by $(\mathbf{a} \mathbf{b} \mathbf{c})$.

Since interchanging of two rows reverses the sign of the determinant (cf. Sec. 0.3), we have

$$(2) \quad (\mathbf{a} \mathbf{b} \mathbf{c}) = -(\mathbf{b} \mathbf{a} \mathbf{c}), \quad \text{etc.}$$

Interchanging twice, we obtain

$$(3) \quad (\mathbf{a} \mathbf{b} \mathbf{c}) = (\mathbf{b} \mathbf{c} \mathbf{a}) = (\mathbf{c} \mathbf{a} \mathbf{b}).$$

Now, by definition,

$$(\mathbf{a} \mathbf{b} \mathbf{c}) = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}), \quad (\mathbf{c} \mathbf{a} \mathbf{b}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}),$$

and since scalar multiplication is commutative, the last expression is equal to $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$. Therefore,

$$(4) \quad \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}.$$

Furthermore, for any constant k ,

$$(5) \quad (k\mathbf{a} \mathbf{b} \mathbf{c}) = k(\mathbf{a} \mathbf{b} \mathbf{c}).$$

The absolute value of the scalar triple product $(\mathbf{a} \mathbf{b} \mathbf{c})$ has a simple geometrical interpretation. It is equal to the volume of the parallelepiped P with \mathbf{a} , \mathbf{b} , \mathbf{c} as adjacent edges.

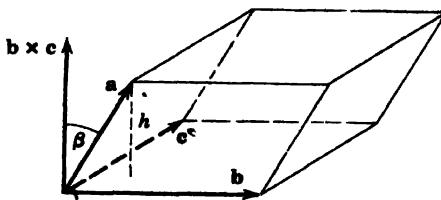


Fig. 142. Geometrical interpretation of a scalar triple product.

Indeed, from (1) in Sec. 5.4 we obtain

$$(a \cdot b \cdot c) = a \cdot (b \times c) = |a| |b \times c| \cos \beta$$

where β is the angle between a and the product vector $b \times c$. Now $|b \times c|$ is the area of the base of P , and the altitude h of P is equal to the absolute value of $|a| \cos \beta$ (Fig. 142). This proves our statement.

From this geometrical consideration it follows that the value of the scalar triple product is a real number which is independent of the choice of right-handed Cartesian coordinates in space. But we should keep in mind that for left-handed Cartesian coordinate systems we have to use (5) instead of (4), Sec. 5.6, and this leads to a minus sign in front of the determinant in (1). We may also say that the value of the determinant is invariant under transformations of right-handed into right-handed, or left-handed into left-handed, Cartesian coordinate systems, but is multiplied by -1 under a transition from a right-handed to a left-handed system (or conversely).

Example 1. Find the volume of the tetrahedron with a, b, c as adjacent edges where, with respect to right-handed Cartesian coordinates,

$$a = i + 2k, \quad b = 4i + 6j + 2k, \quad c = 3i + 3j - 6k.$$

The volume V of the parallelepiped having a, b, c as adjacent edges is obtained from the scalar triple product

$$(a \cdot b \cdot c) = \begin{vmatrix} 1 & 0 & 2 \\ 4 & 6 & 2 \\ 3 & 3 & -6 \end{vmatrix} = -54$$



Fig. 143. Tetrahedron.

and $V = 54$; the minus sign indicates that the vectors a, b, c , in this order, form a left-handed triple. The volume of the tetrahedron is $1/6$ of that of the parallelepiped, namely, 9.

We shall now introduce the basic notions of linear dependence and independence of vectors.

n given vectors $a_{(1)}, \dots, a_{(n)}$ are said to be linearly dependent, if at least one of them can be represented as a "linear combination" of the other $n - 1$ vectors, that is, as the sum of those vectors, each multiplied by a constant (which may be zero or not). If none of the vectors can be represented in this fashion, they are said to be linearly independent.

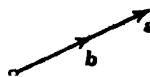


Fig. 144. Collinear vectors.

Consider the vector equation

$$(6) \quad k_1 \mathbf{a}_{(1)} + k_2 \mathbf{a}_{(2)} + \cdots + k_n \mathbf{a}_{(n)} = \mathbf{0},$$

where k_1, \dots, k_n are real numbers. This equation certainly holds for $k_1 = k_2 = \cdots = k_n = 0$. If we can find numbers k_1, \dots, k_n , not all zero, such that (6) holds, then we can divide by a number $k_1 \neq 0$ and represent the corresponding vector $\mathbf{a}_{(1)}$ as a linear combination of the others. (For example, if $k_1 \neq 0$, we obtain from (6)

$$\mathbf{a}_{(1)} = l_2 \mathbf{a}_{(2)} + \cdots + l_n \mathbf{a}_{(n)} \quad \text{where} \quad l_i = -k_i/k_1.)$$

Hence, in this case the vectors are linearly dependent. However, if $k_1 = k_2 = \cdots = k_n = 0$ is the only set of numbers for which (6) holds, we cannot solve (6) for any of the vectors, and the vectors are linearly independent. This proves the following criterion.

n vectors $\mathbf{a}_{(1)}, \dots, \mathbf{a}_{(n)}$ are linearly dependent if, and only if, (6) holds for a set of constants k_1, \dots, k_n , not all zero.

Example 2. The vectors $\mathbf{a} = \mathbf{i} + 2\mathbf{j} + \mathbf{k}$, $\mathbf{b} = 3\mathbf{k}$, $\mathbf{c} = 2\mathbf{i} + 4\mathbf{j}$ are linearly dependent because $6\mathbf{a} - 2\mathbf{b} - 3\mathbf{c} = \mathbf{0}$; we have $\mathbf{a} = \frac{1}{3}\mathbf{b} + \frac{2}{3}\mathbf{c}$. The vectors \mathbf{i} , \mathbf{j} , \mathbf{k} are linearly independent.

If one of those vectors is the zero vector, then the vectors are linearly dependent, because if, say, $\mathbf{a}_{(1)} = \mathbf{0}$, then (6) holds for any number k_1 and $k_2 = \cdots = k_n = 0$.

Two linearly dependent nonzero vectors \mathbf{a} and \mathbf{b} are collinear, that is, if we let their initial points coincide, then they lie in the same line (Fig. 144). In this case, $\mathbf{a} \times \mathbf{b} = \mathbf{0}$. Conversely, if $\mathbf{a} \times \mathbf{b} = \mathbf{0}$, then

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}| |\mathbf{b}| \sin \gamma = 0,$$

and when $\mathbf{a} \neq \mathbf{0}$ and $\mathbf{b} \neq \mathbf{0}$, then $\gamma = 0$ or $\gamma = \pi$. This yields the following result.

Theorem 1. *Two vectors are linearly dependent if, and only if, their vector product is the zero vector.*

Three linearly dependent nonzero vectors \mathbf{a} , \mathbf{b} , and \mathbf{c} are coplanar, that is, if we let their initial points coincide, then they lie in the same plane (Fig. 145). In fact, linear dependence implies that one of the vectors is a linear combination of the other two vectors, for example, $\mathbf{a} = k\mathbf{b} + l\mathbf{c}$, and thus lies in the plane⁴ determined by \mathbf{b} and \mathbf{c} . Remembering the geometrical

⁴ Or in the straight line; if \mathbf{b} and \mathbf{c} are collinear.

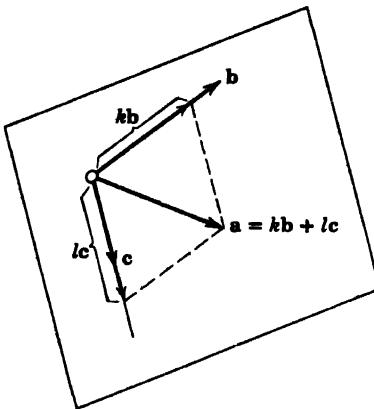


Fig. 145. Coplanar vectors.

interpretation of the scalar triple product, we obtain the following useful criterion.

Theorem 2. *Three vectors are linearly dependent if, and only if, their scalar triple product is zero.*

Example 3. Theorem 2 confirms our results in Ex. 2, because

$$(\mathbf{a} \mathbf{b} \mathbf{c}) = \begin{vmatrix} 1 & 2 & 1 \\ 0 & 0 & 3 \\ 2 & 4 & 0 \end{vmatrix} = 0, \quad (\mathbf{i} \mathbf{j} \mathbf{k}) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1.$$

Finally we have

Theorem 3. *Four or more vectors in three-dimensional space are always linearly dependent.*

Proof. Equation (6) is equivalent to a system of three equations for the three components of the vectors. Since the number of unknowns n (≥ 4) is greater than the number of equations, the system has a solution k_1, \dots, k_n , where not all k_i are zero (Theorem 2, Sec. 7.6), and this means linear dependence of the n vectors.

Problems

Find the value of the scalar triple product of the given vectors (taken in the given order), whose components are referred to right-handed Cartesian coordinates.

- | | |
|---|--|
| 1. $\mathbf{i}, \mathbf{j}, \mathbf{k}$ | 2. $\mathbf{j}, \mathbf{k}, \mathbf{i}$ |
| 3. $\mathbf{i}, \mathbf{k}, \mathbf{j}$ | 4. $2\mathbf{i}, 3\mathbf{j} + \mathbf{k}, 4\mathbf{i} - \mathbf{k}$ |
| 5. $\mathbf{j}, \mathbf{j} + \mathbf{k}, 3\mathbf{i} - 2\mathbf{j}$ | 6. $3\mathbf{i} - \mathbf{j} + \mathbf{k}, \mathbf{j} + \mathbf{k}, -\mathbf{i} + 2\mathbf{j} - 3\mathbf{k}$ |
| 7. $2\mathbf{j} + 3\mathbf{k}, \mathbf{i} + 4\mathbf{j}, 6\mathbf{k}$ | 8. $4\mathbf{i} + 2\mathbf{j} + \mathbf{k}, \mathbf{i} - \mathbf{j} - \mathbf{k}, 3\mathbf{i} + 2\mathbf{j} + 4\mathbf{k}$ |
| 9. $2\mathbf{i} - 3\mathbf{k}, -\mathbf{i}, 3\mathbf{i} - \mathbf{j} + 5\mathbf{k}$ | 10. $\mathbf{i} + \mathbf{j} + \mathbf{k}, 3\mathbf{i}, 4\mathbf{j} + 5\mathbf{k}$ |

Are the following vectors linearly dependent or independent?

11. $3\mathbf{i} - \mathbf{j}$

12. $\mathbf{i} - 6\mathbf{j} + 2\mathbf{k}, 2\mathbf{j} + 7\mathbf{k}, -2\mathbf{i} + 12\mathbf{j} - 4\mathbf{k}$

13. $\mathbf{i} + \mathbf{j}, \mathbf{i} - \mathbf{j}$

14. $\mathbf{i}, \mathbf{0}, \mathbf{j}$

15. $\mathbf{i}, \mathbf{j}, \mathbf{k}, 3\mathbf{i} + 2\mathbf{j} - \mathbf{k}$

16. $4\mathbf{i} + 5\mathbf{j}, \mathbf{i} + 2\mathbf{j}, -\mathbf{i} + 3\mathbf{j}$

17. $\mathbf{i} + \mathbf{j}, \mathbf{j} + \mathbf{k}, \mathbf{k} + \mathbf{i}$

18. $\mathbf{i} + \mathbf{k}, 3\mathbf{i} - 5\mathbf{k}, 8\mathbf{k}$

Find the volume of the parallelepiped that has the given vectors as adjacent edges.

19. $\mathbf{i}, -\mathbf{j}, 3\mathbf{k}$

20. $2\mathbf{j} + \mathbf{k}, \mathbf{i} - \mathbf{j}, -\mathbf{j} + 4\mathbf{k}$

21. $\mathbf{i} + \mathbf{j}, \mathbf{i} - \mathbf{j}, -\mathbf{k}$

22. $\mathbf{i}, 2\mathbf{i} + 4\mathbf{j} - \mathbf{k}, -\mathbf{i} + \mathbf{j} + 3\mathbf{k}$

23. $3\mathbf{i} + 2\mathbf{k}, -4\mathbf{j}, \mathbf{i} + \mathbf{j} - \mathbf{k}$

24. $\mathbf{j} + \mathbf{k}, \mathbf{i} + \mathbf{j}, \mathbf{i} - 3\mathbf{k}$

Find the volume of the tetrahedron that has the following vertices.

25. $(0, 1, 1), (1, 0, 0), (2, 2, 3), (-1, 0, 4)$

26. $(0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1)$

27. $(2, 1, 8), (3, 2, 9), (2, 1, 4), (3, 3, 10)$

28. $(-1, 0, 1), (4, 4, 5), (0, 1, 0), (2, 2, 0)$

29. $(4, 7, -1), (3, 3, 3), (1, -1, 2), (2, 1, 0)$

30. A system of three homogeneous linear equations in three unknowns has a non-trivial solution if, and only if, the determinant of the coefficients of the system is zero. Using this familiar theorem, prove Theorem 2.

5.8 Other Repeated Products

Other types of repeated products that occur frequently in applications may be expressed in terms of scalar, vector, and scalar triple products.

We first show that

(1) $\mathbf{b} \times (\mathbf{c} \times \mathbf{d}) = (\mathbf{b} \cdot \mathbf{d})\mathbf{c} - (\mathbf{b} \cdot \mathbf{c})\mathbf{d}$.

Proof. We choose a right-handed Cartesian coordinate system such that the x -axis has the direction of \mathbf{d} and the xy -plane contains \mathbf{c} . Then the vectors in (1) are of the form

$$\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}, \quad \mathbf{c} = c_1\mathbf{i} + c_2\mathbf{j}, \quad \mathbf{d} = d_1\mathbf{i}.$$

Hence, $\mathbf{c} \times \mathbf{d} = -c_2d_1\mathbf{k}$ and, furthermore,

$$\mathbf{b} \times (\mathbf{c} \times \mathbf{d}) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ b_1 & b_2 & b_3 \\ 0 & 0 & -c_2d_1 \end{vmatrix} = -b_2c_2d_1\mathbf{i} + b_1c_2d_1\mathbf{j}.$$

On the other hand, we obtain

$$(\mathbf{b} \cdot \mathbf{d})\mathbf{c} - (\mathbf{b} \cdot \mathbf{c})\mathbf{d} = b_1d_1(c_1\mathbf{i} + c_2\mathbf{j}) - (b_1c_1 + b_2c_2)d_1\mathbf{i} = b_1c_2d_1\mathbf{j} - b_2c_2d_1\mathbf{i}.$$

This proves (1) for our special coordinate system. Now the magnitude and direction of a vector and a vector product, and the value of a scalar product,

are independent of the choice of the coordinates. Furthermore, the representation of $\mathbf{b} \times (\mathbf{c} \times \mathbf{d})$ in terms of $\mathbf{i}, \mathbf{j}, \mathbf{k}$ will be the same for right-handed and left-handed systems, because of the double cross-multiplication. Hence, (1) holds in any Cartesian coordinate system, and the proof is complete.

From (1) we obtain

$$(\mathbf{b} \times \mathbf{c}) \times \mathbf{d} = -\mathbf{d} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{d} \cdot \mathbf{b})\mathbf{c} - (\mathbf{d} \cdot \mathbf{c})\mathbf{b}.$$

This shows that in general $\mathbf{b} \times (\mathbf{c} \times \mathbf{d})$ and $(\mathbf{b} \times \mathbf{c}) \times \mathbf{d}$ are different vectors, that is, *cross-multiplication is not associative*, and the parentheses in (1) are important and cannot be omitted.

Example 1. With respect to a right-handed Cartesian coordinate system we have

$$(\mathbf{i} \times \mathbf{j}) \times \mathbf{j} = \mathbf{k} \times \mathbf{j} = -\mathbf{i}, \quad \text{but} \quad \mathbf{i} \times (\mathbf{j} \times \mathbf{j}) = \mathbf{0}.$$

Taking the scalar product of both sides of (1) and any vector \mathbf{a} , we find

$$\mathbf{a} \cdot [\mathbf{b} \times (\mathbf{c} \times \mathbf{d})] = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}).$$

On the other hand, by using scalar triple products, we have [cf. (4) in the last section]

$$\mathbf{a} \cdot [\mathbf{b} \times (\mathbf{c} \times \mathbf{d})] = (\mathbf{a} \cdot \mathbf{b})(\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}).$$

By comparing these two results we obtain the important identity of Lagrange

$$(2) \quad (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}).$$

If we replace \mathbf{b} by $\mathbf{a} \times \mathbf{b}$ in (1), then we obtain

$$(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) = [(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{d}]\mathbf{c} - [(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}]\mathbf{d}$$

and, therefore,

$$(3) \quad (\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{b})(\mathbf{d} \cdot \mathbf{c}) - (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}).$$

Problems

With respect to a right-handed Cartesian coordinate system, let $\mathbf{a} = \mathbf{i} + 2\mathbf{j} - \mathbf{k}$, $\mathbf{b} = 3\mathbf{i} - 4\mathbf{k}$, $\mathbf{c} = -\mathbf{i} + \mathbf{j}$, and $\mathbf{d} = 2\mathbf{i} - \mathbf{j} + 3\mathbf{k}$. Find

1. $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$, $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$
 2. $(\mathbf{b} \times \mathbf{c}) \times \mathbf{d}$, $\mathbf{d} \times (\mathbf{b} \times \mathbf{c})$
 3. $(\mathbf{a} \times \mathbf{c}) \times \mathbf{d}$, $(\mathbf{a} \times \mathbf{d}) \times \mathbf{c}$
 4. $(\mathbf{a} \times \mathbf{c}) \times \mathbf{c}$, $(\mathbf{c} \times \mathbf{a}) \times \mathbf{c}$
 5. $(\mathbf{c} \times \mathbf{b}) \times \mathbf{a}$, $[(\mathbf{c} \times \mathbf{b}) \times \mathbf{a}]$
 6. $(\mathbf{b} \times \mathbf{b}) \times \mathbf{c}$, $\mathbf{b} \times (\mathbf{b} \times \mathbf{c})$
 7. $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d})$
 8. $(\mathbf{a} \times \mathbf{c}) \cdot (\mathbf{b} \times \mathbf{d})$, $(\mathbf{c} \times \mathbf{a}) \cdot (\mathbf{d} \times \mathbf{b})$
 9. $(\mathbf{a} - \mathbf{b}) \times (\mathbf{c} \times \mathbf{d})$
 10. $(2\mathbf{a} + 3\mathbf{b}) \times \mathbf{c}$, $(\mathbf{a} + \mathbf{b}) \times 6\mathbf{c}$
 11. $(2\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$, $(\mathbf{a} \times \mathbf{b}) \times 2\mathbf{c}$
 12. $(\mathbf{b} \times \mathbf{c}) \cdot (\mathbf{b} \times \mathbf{d})$
 13. $(\mathbf{a} \cdot \mathbf{b})\mathbf{c} \times \mathbf{d}$
 14. $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$
 15. $(\mathbf{b} \times \mathbf{a}) \times (\mathbf{c} \times \mathbf{d})$
 16. $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d})$
17. Let the vectors \mathbf{a} and \mathbf{b} determine a plane P , and let \mathbf{c} be directed along a line L . What does it mean geometrically if $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = 0$?
18. What does it mean geometrically if in Prob. 17, $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = 0$?

19. Let the pairs a, b and c, d each determine a plane. Show that these planes are parallel if, and only if, $(a \times b) \times (c \times d) = 0$.
20. Show that if the vector $v = (a \times b) \times (c \times d)$ is not the null vector, it is directed along the line of intersection of the plane of a and b and the plane of c and d .

5.9 Scalar Fields and Vector Fields

¶ A scalar function is a single-valued function which is defined at certain points in space and whose values are real numbers depending only on the points in space but not on the particular choice of the coordinate system. In most applications the domain of definition D of a scalar function f will be a curve, a surface, or a three-dimensional region in space. The function f associates with each point in D a scalar, a real number, and we say that a **scalar field** is given in D .

If we introduce coordinates x, y, z , then f may be represented in terms of the coordinates, and we write $f(x, y, z)$, but we should keep in mind that the value of f at any point P is independent of the particular choice of coordinates. In order to indicate this fact it is also customary to write $f(P)$ instead of $f(x, y, z)$. The function f may also depend on parameters such as time.

Example 1. The distance $f(P)$ of any point P from a fixed point P_0 in space is a scalar function whose domain of definition D is the whole space. $f(P)$ defines a scalar field in space. If we introduce a Cartesian coordinate system and P_0 has the coordinates x_0, y_0, z_0 , then f is given by the well-known formula

$$f(P) = f(x, y, z) = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$$

where x, y, z are the coordinates of P . If we replace the given Cartesian coordinate system by another such system, then the values of the coordinates of P and P_0 will in general change, but $f(P)$ will have the same value as before. Hence, $f(P)$ is a scalar function. The direction cosines of the line through P and P_0 are not scalars because their values will depend on the choice of the coordinate system.

Example 2. The temperature T within a body B is a scalar function. It defines a scalar field, namely, the temperature field in B . The function T may depend on the time or other parameters. Other examples of scalar fields are the pressure within a region through which a compressible fluid is flowing, and the density of the air of the earth's atmosphere.

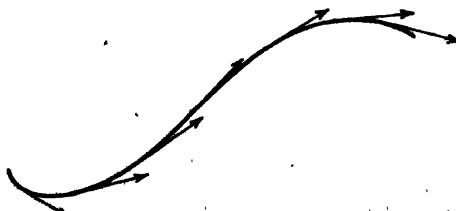


Fig. 146. Field of tangent vectors of a curve.

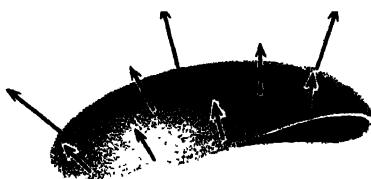


Fig. 147. Field of normal vectors of a surface.

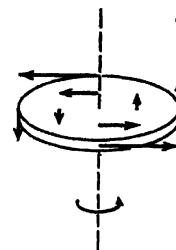


Fig. 148. Velocity field of a rotating body.

If to each point P of a certain set of points in space (for example, the points of a curve, a surface, or a three-dimensional region) a vector $v(P)$ is assigned, then a vector field is said to be given at those points, and $v(P)$ is called a vector function.

If we introduce Cartesian coordinates x, y, z , then we may write

$$v(P) = v(x, y, z) = v_1(x, y, z)i + v_2(x, y, z)j + v_3(x, y, z)k,$$

but we should keep in mind that v depends only on the points of its domain of definition, and at any such point defines the same vector for every choice of the coordinate system.

Example 3 (Velocity field). At any instant the velocity vectors $v(P)$ of a rotating body B constitute a vector field, the so-called *velocity field* of the rotation. If we introduce a Cartesian coordinate system having the origin on the axis of rotation, then (cf. Ex. 3 in Sec. 5.6)

$$(1) \quad v(P) = v(x, y, z) = w \times (xi + yj + zk)$$

where x, y, z are the coordinates of any point P of B at the instant under consideration. If the coordinates are such that the z -axis is the axis of rotation and w points in the positive z -direction, then $w = \omega k$ and

$$v = \begin{vmatrix} i & j & k \\ 0 & 0 & \omega \\ x & y & z \end{vmatrix} = \omega(-yi + xj).$$

Example 4 (Field of force). Let a particle A of mass M be fixed at a point P_0 and let a particle B of mass m be free to take up various positions P in space. Then A attracts B . According to Newton's law of gravitation the corresponding gravitational force p is directed from P to P_0 , and its magnitude is proportional to $1/r^2$ where r is the distance between P and P_0 , say,

$$(2) \quad |p| = \frac{c}{r^2}, \quad c = GMm,$$

where $G (= 6.67 \cdot 10^{-8} \text{ cm}^3/\text{gm} \cdot \text{sec}^2)$ is the gravitational constant. Hence p defines a vector field in space. If we introduce Cartesian coordinates such

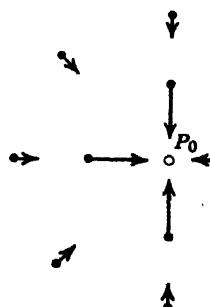


Fig. 149. Gravitational field.

that P_0 has the coordinates x_0, y_0, z_0 and P has the coordinates x, y, z , then

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} \quad (\geq 0).$$

Assuming that $r > 0$ and introducing the vector

$$\mathbf{r} = (x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k},$$

we have $|\mathbf{r}| = r$, and $-\mathbf{r}/r$ is a unit vector in the direction of \mathbf{p} ; the minus sign indicates that \mathbf{p} is directed from P to P_0 (Fig. 149). From this and (2) we obtain

$$(3) \quad \mathbf{p} = |\mathbf{p}|\left(-\frac{\mathbf{r}}{r}\right) = -c \frac{\mathbf{r}}{r^3} = -c \frac{x - x_0}{r^3} \mathbf{i} - c \frac{y - y_0}{r^3} \mathbf{j} - c \frac{z - z_0}{r^3} \mathbf{k}.$$

This vector function describes the gravitational force acting on B .

Problems

Determine the isobars (curves of constant pressure p) of the pressure fields in the xy -plane given by the following functions. Plot some of the isobars.

1. $p = y$

2. $p = xy$

3. $p = x^2 - y^2$

4. $p = \ln(x^2 + y^2)$

5. $p = \arctan \frac{y}{x}$

6. $p = e^x \cos y$

Find the isotherms (curves of constant temperature T) of the following temperature fields in the xy -plane.

7. $T = x + y$

8. $T = x - y$

9. $T = 3x^2y - y^3$

Consider Prob. 9, as follows.

10. Find the temperature at the points $(1, 1)$, $(-2, 2)$, and $(3, 4)$.

11. Find the region in which $T > 0$.

12. Find and graph T on the lines $y = x$ and $y = -x$.

13. Find a formula for T on the unit circle $x^2 + y^2 = 1$.

14. Draw a figure of $T(x, y)$ as a surface in space.

Find the level surfaces (surfaces $f = \text{const}$) of the scalar fields in space given by the following functions.

15. $f = x + y + z$

16. $f = x + y$

17. $f = x^2 + y^2 + z^2$

18. $f = (x^2 + y^2 + z^2)^{-1}$

19. $f = x^2 + y^2 - z$

20. $f = x^2 + 2y^2 + 4z^2$

Consider Prob. 19, as follows.

21. Find the values of f at the points $(0, 0, 0)$, $(1, -1, 3)$, and $(2, 0, 4)$.

22. Find the region in the plane $z = 9$ in which $f < 0$.

23. Find and graph the level surface $f = 0$ in space.

24. Find and graph the level curves (curves $f = \text{const}$) in the plane $z = 4$.

25. Find and graph the level curves of f in the coordinate planes.

Draw figures (similar to Fig. 149) of the vector fields in the xy -plane given by the following vector functions v .

26. $v = 2\mathbf{i}$

27. $v = 3\mathbf{i} - 2\mathbf{j}$

28. $v = x\mathbf{i} + y\mathbf{j}$

29. $v = y\mathbf{i} + x\mathbf{j}$

30. $v = 6xy\mathbf{i} + 3(x^2 - y^2)\mathbf{j}$

31. $v = 2xi - 2yj$

Consider Prob. 30, as follows.

- 32. Find \mathbf{v} at the points $(0, 0)$, $(1, 2)$, and $(-2, 3)$.
- 33. Find and graph \mathbf{v} on the lines $y = x$ and $y = -x$.
- 34. Find the curves on which $|\mathbf{v}| = \text{const.}$
- 35. Show that $\mathbf{v} = (\partial T/\partial x)\mathbf{i} + (\partial T/\partial y)\mathbf{j}$ where T is given in Prob. 9.

Find and graph the curves on which \mathbf{v} has constant length and the curves on which \mathbf{v} has constant direction, where

$$36. \mathbf{v} = xi + 2yj$$

$$38. \mathbf{v} = yi + x^2j$$

$$37. \mathbf{v} = (x^2 - y^2)\mathbf{i} + 2xy\mathbf{j}$$

$$39. \mathbf{v} = (x - y)\mathbf{i} + (x + y)\mathbf{j}$$

- 40. Find the surfaces on which $\mathbf{v} = 3xi + 2zj + yk$ has constant length.

5.10 Vector Calculus

The basic concepts of calculus, such as convergence, continuity, and differentiability, can be introduced to vector analysis in a simple and natural way as follows.

An infinite sequence of vectors $\mathbf{a}_{(n)}$, $n = 1, 2, \dots$, is said to converge to a limit vector \mathbf{a} if

$$(1) \quad \lim_{n \rightarrow \infty} |\mathbf{a}_{(n)} - \mathbf{a}| = 0,$$

and then we write

$$(2) \quad \lim_{n \rightarrow \infty} \mathbf{a}_{(n)} = \mathbf{a}.$$

Clearly, if a Cartesian coordinate system has been introduced, then that sequence of vectors converges to \mathbf{a} if, and only if, the three sequences of the components of the vectors converge to the corresponding components of \mathbf{a} . The simple proof is left to the reader.

Similarly, a vector function $\mathbf{u}(t)$ of a real variable t is said to have the limit \mathbf{l} as t approaches t_0 , if $\mathbf{u}(t)$ is defined in some neighborhood^b of t_0 (possibly except at t_0) and

$$(3) \quad \lim_{t \rightarrow t_0} |\mathbf{u}(t) - \mathbf{l}| = 0.$$

Then we write

$$(4) \quad \lim_{t \rightarrow t_0} \mathbf{u}(t) = \mathbf{l}.$$

A vector function $\mathbf{u}(t)$ is said to be continuous at $t = t_0$ if it is defined in some neighborhood of t_0 and

$$(5) \quad \lim_{t \rightarrow t_0} \mathbf{u}(t) = \mathbf{u}(t_0).$$

^b That is, in some interval on the t -axis containing t_0 as an interior point.

If a Cartesian coordinate system has been introduced, then $\mathbf{u}(t)$ is continuous at t_0 if, and only if, its three components are continuous at t_0 .

A vector function $\mathbf{u}(t)$ is said to be *differentiable* at a point t if the limit

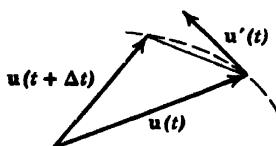
$$(6) \quad \mathbf{u}'(t) = \lim_{\Delta t \rightarrow 0} \frac{\mathbf{u}(t + \Delta t) - \mathbf{u}(t)}{\Delta t}$$

exists. The vector $\mathbf{u}'(t)$ is called the **derivative** of $\mathbf{u}(t)$. Cf. Fig. 150.

With respect to any Cartesian coordinate system, let

$$\mathbf{u}(t) = u_1(t)\mathbf{i} + u_2(t)\mathbf{j} + u_3(t)\mathbf{k}.$$

Then $\mathbf{u}'(t)$ exists if, and only if, the derivatives



$$u_m'(t) = \lim_{\Delta t \rightarrow 0} \frac{u_m(t + \Delta t) - u_m(t)}{\Delta t}, \quad m = 1, 2, 3,$$

of the three components exist, and

$$(7) \quad \mathbf{u}'(t) = u_1'(t)\mathbf{i} + u_2'(t)\mathbf{j} + u_3'(t)\mathbf{k};$$

that is, to differentiate a vector function one differentiates each component separately.

The familiar rules of differentiation yield corresponding rules for differentiating vector functions, for example,

$$(cu)' = cu' \quad (c \text{ constant}), \quad (\mathbf{u} + \mathbf{v})' = \mathbf{u}' + \mathbf{v}'$$

and in particular

$$(8) \quad (\mathbf{u} \cdot \mathbf{v})' = \mathbf{u}' \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{v}'$$

$$(9) \quad (\mathbf{u} \times \mathbf{v})' = \mathbf{u}' \times \mathbf{v} + \mathbf{u} \times \mathbf{v}'$$

$$(10) \quad (\mathbf{u} \cdot \mathbf{v} \cdot \mathbf{w})' = (\mathbf{u}' \cdot \mathbf{v} \cdot \mathbf{w}) + (\mathbf{u} \cdot \mathbf{v}' \cdot \mathbf{w}) + (\mathbf{u} \cdot \mathbf{v} \cdot \mathbf{w}').$$

The simple proofs are left to the reader. In (9), the order of the vectors must be carefully observed because cross-multiplication is not commutative.

Example 1. Let $\mathbf{u}(t)$ be a vector function whose magnitude is constant, say, $|\mathbf{u}(t)| = c$. Then $|\mathbf{u}|^2 = \mathbf{u} \cdot \mathbf{u} = c^2$, and $(\mathbf{u} \cdot \mathbf{u})' = 2\mathbf{u} \cdot \mathbf{u}' = 0$, by differentiation [cf. (8)]. This yields the following result. *The derivative of a vector function $\mathbf{u}(t)$ of constant magnitude is either the zero vector or is orthogonal to $\mathbf{u}(t)$.*

Important applications of derivatives will be considered in the following sections.

From our considerations the way of introducing partial differentiation to vector analysis is obvious. Let the components of a vector function

$$\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$$

be differentiable functions of n variables t_1, \dots, t_n . Then the partial derivative of \mathbf{u} with respect to t_i is denoted by $\partial \mathbf{u} / \partial t_i$ and is defined as the

vector function

$$\frac{\partial \mathbf{u}}{\partial t_1} = \frac{\partial u_1}{\partial t_1} \mathbf{i} + \frac{\partial u_2}{\partial t_1} \mathbf{j} + \frac{\partial u_3}{\partial t_1} \mathbf{k}.$$

Similarly,

$$\frac{\partial^2 \mathbf{u}}{\partial t_1 \partial t_m} = \frac{\partial^2 u_1}{\partial t_1 \partial t_m} \mathbf{i} + \frac{\partial^2 u_2}{\partial t_1 \partial t_m} \mathbf{j} + \frac{\partial^2 u_3}{\partial t_1 \partial t_m} \mathbf{k},$$

and so on.

Example 2. Let $\mathbf{r}(t_1, t_2) = a \cos t_1 \mathbf{i} + a \sin t_1 \mathbf{j} + t_2 \mathbf{k}$. Then

$$\frac{\partial \mathbf{r}}{\partial t_1} = -a \sin t_1 \mathbf{i} + a \cos t_1 \mathbf{j}, \quad \frac{\partial \mathbf{r}}{\partial t_2} = \mathbf{k}.$$

Note that if $\mathbf{r}(t_1, t_2)$ is interpreted as a position vector, it represents a cylinder of revolution of radius a , having the z -axis as axis of rotation. (Representations of surfaces will be considered in Sec. 6.5.)

Problems

1. Prove (8). (Use (8) in Sec. 5.4.)
2. Prove (9).
3. Derive (10) from (8) and (9).
4. Prove (10) directly.
5. Show that if $\mathbf{u}(t)$ is a unit vector and $\mathbf{u}'(t) \neq 0$, then \mathbf{u} and \mathbf{u}' are orthogonal.

Find \mathbf{u}' , $|\mathbf{u}'|$, \mathbf{u}'' , and $|\mathbf{u}''|$ where

6. $\mathbf{u} = \mathbf{a} + b\mathbf{t}$

7. $\mathbf{u} = t\mathbf{i} + t^2\mathbf{j}$

8. $\mathbf{u} = \cos t \mathbf{i} + \sin t \mathbf{j}$

9. $\mathbf{u} = \cos t \mathbf{i} + 4 \sin t \mathbf{j} + \mathbf{k}$

10. $\mathbf{u} = t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k}$

11. $\mathbf{u} = 2 \cos t \mathbf{i} + 2 \sin t \mathbf{j} + t\mathbf{k}$

Let $\mathbf{u} = t\mathbf{i} + 2t^2\mathbf{k}$, $\mathbf{v} = t^3\mathbf{j} + t\mathbf{k}$, and $\mathbf{w} = \mathbf{i} + t\mathbf{j} + t^2\mathbf{k}$. Find

12. $(\mathbf{u} \cdot \mathbf{v})'$

13. $(\mathbf{u} \times \mathbf{v})'$

14. $(\mathbf{u} \cdot \mathbf{w})'$

15. $[(\mathbf{u} \times \mathbf{v}) \times \mathbf{w}]'$

16. $[\mathbf{u} \times (\mathbf{v} \times \mathbf{w})]'$

17. $[(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w}]'$

Find the first partial derivatives of the following vector functions.

18. $\mathbf{v} = xi + 2y\mathbf{j}$

19. $\mathbf{v} = (x^2 - y^2)\mathbf{i} + 2xy\mathbf{j}$

20. $\mathbf{v} = x^2\mathbf{i} + y^2\mathbf{j} + z^2\mathbf{k}$

21. $\mathbf{v} = yzi + zx\mathbf{j} + xy\mathbf{k}$

22. Show that the equation $\mathbf{u}'(t) = \mathbf{c}$ has the solution $\mathbf{u}(t) = ct + \mathbf{b}$ where \mathbf{c} and \mathbf{b} are constant vectors.

23. Show that $\mathbf{u}(t) = b e^{\lambda t} + c e^{-\lambda t}$ satisfies the equation $\mathbf{u}'' - \lambda^2 \mathbf{u} = 0$. (\mathbf{b} and \mathbf{c} are constant vectors.)

24. Since $\mathbf{u} \times \mathbf{u} = \mathbf{0}$, also $(\mathbf{u} \times \mathbf{u})' = \mathbf{0}$. How does this follow from (9)?

25. Find formulas similar to (8) and (9) for $(\mathbf{u} \cdot \mathbf{v})''$ and $(\mathbf{u} \times \mathbf{v})''$.

5.11 Curves

As an important application of vector calculus, let us now consider some basic facts about curves in space. The student will know that curves occur in many considerations in calculus as well as in physics, for example, as

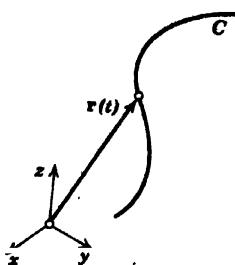


Fig. 151. Parametric representation of a curve.

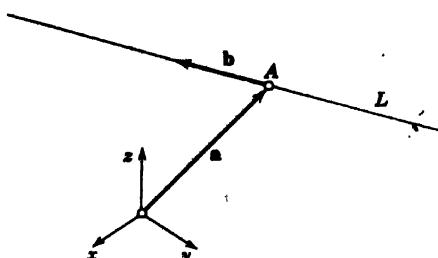


Fig. 152. Parametric representation of a straight line.

paths of moving particles. The consideration will be a part of an important branch of mathematics, which is called **differential geometry** and which may be defined as the study of curves and surfaces by means of calculus. Cf. Ref. [C4] in Appendix 1.

A Cartesian coordinate system being given, we may represent a curve C by a vector function (Fig. 151)

$$(1) \quad \mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k};$$

to each value t_0 of the real variable t there corresponds a point of C having the position vector $\mathbf{r}(t_0)$, that is, the coordinates $x(t_0)$, $y(t_0)$, $z(t_0)$.

A representation of the form (1) is called a **parametric representation** of the curve C , and t is called the **parameter** of this representation. This type of representation is useful in many applications, for example, in mechanics where the variable t may be the time.

Other types of representations of curves in space are

$$(2) \quad y = f(x), \quad z = g(x)$$

and

$$(3) \quad F(x, y, z) = 0, \quad G(x, y, z) = 0.$$

By setting $x = t$ we may write (2) in the form (1), namely

$$\mathbf{r}(t) = t\mathbf{i} + f(t)\mathbf{j} + g(t)\mathbf{k}.$$

In (3), each equation represents a surface, and the curve is the intersection of the two surfaces.

A **plane curve** is a curve which lies in a plane in space. A curve which is not a plane curve is called a **twisted curve**.

Example 1 (Straight line). Any straight line L can be represented in the form

$$(4) \quad \mathbf{r}(t) = \mathbf{a} + t\mathbf{b} = (a_1 + tb_1)\mathbf{i} + (a_2 + tb_2)\mathbf{j} + (a_3 + tb_3)\mathbf{k}$$

where \mathbf{a} and \mathbf{b} are constant vectors. L passes through the point A with position vector $\mathbf{r} = \mathbf{a}$ and has the direction of \mathbf{b} (Fig. 152). If \mathbf{b} is a unit vector, its components are the **direction cosines** of L , and in this case, $|t|$ measures the distance of the points of L from A .

Example 2 (Ellipse, circle). The vector function

$$(5) \quad \mathbf{r}(t) = a \cos t \mathbf{i} + b \sin t \mathbf{j}$$

represents an ellipse in the xy -plane with center at the origin and principal axes in the direction of the x and y axes. In fact, since $\cos^2 t + \sin^2 t = 1$, we obtain from (5)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad z = 0.$$

If $b = a$, then (5) represents a *circle* of radius a .

Example 3 (Circular helix). The twisted curve C represented by the vector function

$$(6) \quad \mathbf{r}(t) = a \cos t \mathbf{i} + a \sin t \mathbf{j} + ct \mathbf{k} \quad (c \neq 0)$$

is called a *circular helix*. It lies on the cylinder $x^2 + y^2 = a^2$. If $c > 0$, the helix is shaped like a right-handed screw (Fig. 153). If $c < 0$, it looks like a left-handed screw (Fig. 154).

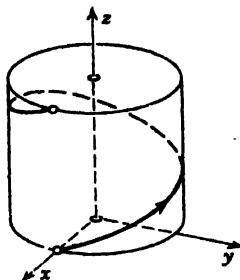


Fig. 153. Right-handed circular helix.

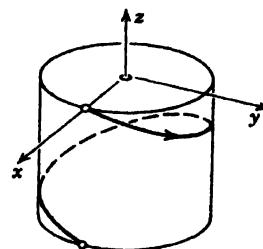


Fig. 154. Left-handed circular helix.

The portion between any two points of a curve is often called an *arc of a curve*. For the sake of simplicity we shall use the single term "curve" to denote an entire curve as well as an arc of a curve.

A curve may have self-intersections; the points of intersection are called *multiple points* of the curve. A curve having no multiple points is called a *simple curve*.



Fig. 155. Curves having double points.

Example 4. Ellipses and helices are simple curves. The curve represented by

$$\mathbf{r}(t) = (t^2 - 1)\mathbf{i} + (t^3 - t)\mathbf{j}$$

is not simple since it has a double point at the origin; this point corresponds to the two values $t = 1$ and $t = -1$.

We finally mention that a given curve C may be represented by various vector functions. For example, if C is represented by (1) and we set $t = h(t^*)$, then we obtain a new vector function $\mathbf{r}(t^*)$ representing C , provided $h(t^*)$ takes on all the values of t occurring in (1).

Example 5. The parabola $y = x^2$ in the xy -plane may be represented by the vector function

$$\mathbf{r}(t) = t\mathbf{i} + t^2\mathbf{j} \quad (-\infty < t < \infty).$$

If we set $t = -2t^*$, we obtain another representation of the parabola:

$$\mathbf{r}(t^*) = -2t^*\mathbf{i} + 4t^{*2}\mathbf{j}.$$

If we set $t = t^{*2}$, we obtain

$$\mathbf{r}(t^*) = t^{*2}\mathbf{i} + t^{*4}\mathbf{j},$$

but this function represents only the portion of the parabola in the first quadrant, because $t^{*2} \geq 0$ for all t^* .

Problems

Find a parametric representation of the straight line through a point A in the direction of a vector \mathbf{b} , where

1. $A: (0, 0, 0), \mathbf{b} = \mathbf{i} + \mathbf{j}$

2. $A: (1, 3, 2), \mathbf{b} = -\mathbf{i} + \mathbf{k}$

3. $A: (2, 1, 0), \mathbf{b} = 2\mathbf{j} + \mathbf{k}$

4. $A: (0, 4, 1), \mathbf{b} = \mathbf{i} - \mathbf{j} + 2\mathbf{k}$

Find a parametric representation of the straight line through the points A and B , where

5. $A: (0, 0, 0), B: (1, 1, 1)$

6. $A: (-1, 8, 3), B: (1, 0, 0)$

7. $A: (1, 5, 3), B: (0, 2, -1)$

8. $A: (1, 4, 2), B: (1, 4, -2)$

Find a parametric representation of the straight line represented by

9. $y = x, z = 0$

10. $7x - 3y + z = 14, 4x - 3y - 2z = -1$

11. $x + y = 0, x - z = 0$

12. $x + y + z = 1, y - z = 0$

Represent the following curves in parametric form.

13. $x^2 + y^2 = 1, z = 0$

14. $y = x^4, z = 0$

15. $y = x^2, z = x^3$

16. $x^2 + y^2 - 2x - 4y = -1, z = 0$

17. $4(x + 1)^2 + y^2 = 4, z = 0$

18. $x^2 + y^2 = 4, z = e^x$

19. Determine the orthogonal projections of the circular helix (6) in the coordinate planes.

Sketch figures of the curves represented by the following functions $\mathbf{r}(t)$.

20. $t\mathbf{i} + 2t\mathbf{j} - t\mathbf{k}$

21. $2 \sin t \mathbf{i} + 2 \cos t \mathbf{j}$

22. $(1 + \cos t)\mathbf{i} + \sin t \mathbf{j}$

23. $\cos t \mathbf{i} + 2 \sin t \mathbf{j}$

24. $\cos t \mathbf{i} + \sin t \mathbf{j} + t\mathbf{k}$

25. $t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k}$

5.12 Arc Length

To define the length of a curve C we may proceed as follows. We inscribe in C a broken line of n chords joining the two end points of C as shown in Fig. 156. This we do for each positive integer n in an arbitrary way but such that the maximum chord-length approaches zero as n approaches infinity. The lengths of these lines of chords can be obtained from the theorem of Pythagoras. If the sequence of these lengths l_1, l_2, \dots is convergent with the limit l , then C is said to be rectifiable, and l is called the length of C .

If C can be represented by a continuously differentiable⁶ vector function

$$\mathbf{r} = \mathbf{r}(t) \quad (a \leq t \leq b),$$

then it can be shown that C is rectifiable, and its length l is given by the integral

$$(1) \quad l = \int_a^b \sqrt{\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}} dt \quad \left(\dot{\mathbf{r}} = \frac{d\mathbf{r}}{dt} \right),$$

whose value is independent of the choice of the parametric representation. The proof is quite similar to that for plane curves usually considered in elementary integral calculus (cf. Ref. [A13]) and can be found in Ref. [C4] in Appendix 1.

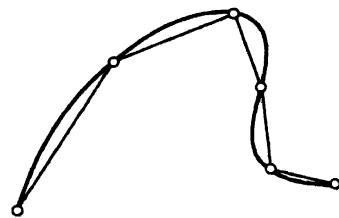


Fig. 156. Length of a curve.

If we replace the fixed upper limit b in (1) with a variable upper limit t , the integral becomes a function of t , say, $s(t)$; denoting the variable of integration by t^* , we have

$$(2) \quad s(t) = \int_a^t \sqrt{\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}} dt^* \quad \left(\dot{\mathbf{r}} = \frac{d\mathbf{r}}{dt^*} \right).$$

This function $s(t)$ is called the **arc length** of C .

The arc length s may serve as a parameter in parametric representations of curves, and we shall see that this will lead to a simplification of various formulas.

The constant a in (2) may be replaced by another constant; that is, the point of the curve corresponding to $s = 0$ may be chosen in an arbitrary manner. The sense corresponding to increasing values of s is called the **positive sense** on C ; in this fashion any representation $\mathbf{r}(s)$ or $\mathbf{r}(t)$ of C defines a certain **orientation** of C . Obviously, there are two ways of *orienting* C , and it is not difficult to see that the transition from one orientation to the opposite orientation can be effected by a transformation of the parameter whose derivative is negative

From (2) we obtain

$$(3) \quad \left(\frac{ds}{dt} \right)^2 = \frac{d\mathbf{r}}{dt} \cdot \frac{d\mathbf{r}}{dt} = \left(\frac{dx}{dt} \right)^2 + \left(\frac{dy}{dt} \right)^2 + \left(\frac{dz}{dt} \right)^2.$$

It is customary to write

$$d\mathbf{r} = dx\mathbf{i} + dy\mathbf{j} + dz\mathbf{k}$$

and

$$(4) \quad ds^2 = d\mathbf{r} \cdot d\mathbf{r} = dx^2 + dy^2 + dz^2.$$

ds is called the **linear element** of C .

⁶ "Continuously differentiable" means that the derivative exists and is continuous; "twice continuously differentiable" means that the first and second derivatives exist and are continuous, and so on.

Example 1. In the case of the circle

$$\mathbf{r}(t) = a \cos t \mathbf{i} + a \sin t \mathbf{j}$$

we have $\dot{\mathbf{r}} = -a \sin t \mathbf{i} + a \cos t \mathbf{j}$, $\dot{\mathbf{r}} \cdot \dot{\mathbf{r}} = a^2$, and therefore,

$$s(t) = \int_0^t a dt^* = at.$$

Hence a representation of the circle with the arc length as parameter is

$$\mathbf{r}(s) = a \cos \frac{s}{a} \mathbf{i} + a \sin \frac{s}{a} \mathbf{j}.$$

The circle is oriented in the counterclockwise sense, which corresponds to increasing values of s . Setting $s = -\tilde{s}$, we obtain

$$\tilde{\mathbf{r}}(\tilde{s}) = a \cos \frac{\tilde{s}}{a} \mathbf{i} - a \sin \frac{\tilde{s}}{a} \mathbf{j};$$

we have $ds/d\tilde{s} = -1 < 0$, and the circle is now oriented in the clockwise sense. Note that we wrote $\tilde{\mathbf{r}}$ because this is a new function; in fact, $\mathbf{r}(-\tilde{s}) = \tilde{\mathbf{r}}(\tilde{s})$.

Problems

Graph the following curves and find their lengths.

1. *Catenary* $y = \cosh x$, $z = 0$, from $x = 0$ to $x = 1$ (a, 0, 2ae)
2. *Circular helix* $\mathbf{r}(t) = a \cos t \mathbf{i} + a \sin t \mathbf{j} + ct \mathbf{k}$ from $(a, 0, 0)$ to ~~(a, 0, 2ae)~~
3. *Semicubical parabola* $y = x^{3/2}$, $z = 0$, from $(0, 0, 0)$ to $(4, 8, 0)$
4. *Four-cusped hypocycloid* $\mathbf{r}(t) = a \cos^3 t \mathbf{i} + a \sin^3 t \mathbf{j}$, total length
5. If a plane curve is represented in the form $y = f(x)$, $z = 0$, show that its length between $x = a$ and $x = b$ is

$$l = \int_a^b \sqrt{1 + y'^2} dx.$$

6. Using the formula in Prob. 5, find the length of a circle of radius a .
7. Show that if a plane curve is represented in polar coordinates $\rho = \sqrt{x^2 + y^2}$ and $\theta = \arctan(y/x)$, then $ds^2 = \rho^2 d\theta^2 + d\rho^2$.
8. Using the formula in Prob. 7, find the length of a circle of radius a .
9. Graph the *cardioid* $\rho = a(1 - \cos \theta)$ and find its length.
10. If a curve is represented by a parametric representation, show that a transformation of the parameter whose derivative is negative reverses the orientation.

5.13 Tangent, Curvature and Torsion

The *tangent* to a curve C at a point P of C is defined as the limiting position of the straight line L through P and another point Q of C as Q approaches P along the curve (Fig. 157).

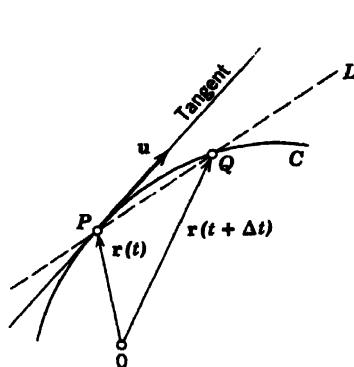


Fig. 157. Tangent to a curve.

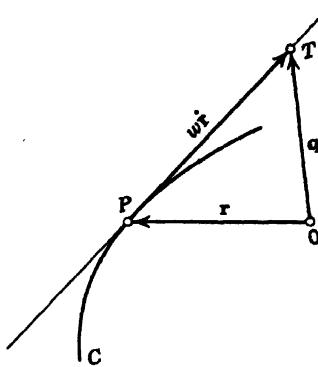


Fig. 158. Representation of the tangent to a curve.

Suppose that C is represented by a continuously differentiable vector function $\mathbf{r}(t)$ where t is any parameter. Let P and Q correspond to t and $t + \Delta t$, respectively. Then L has the direction of the vector

$$[\mathbf{r}(t + \Delta t) - \mathbf{r}(t)]/\Delta t.$$

Hence, if the vector

$$(1) \quad \dot{\mathbf{r}} = \lim_{\Delta t \rightarrow 0} \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t)}{\Delta t}$$

is not the zero vector, it has the direction of the tangent to C at P . It points in the direction of increasing values of t , and its sense, therefore, depends on the orientation of the curve. $\dot{\mathbf{r}}$ is called a *tangent vector* of C at P , and the corresponding unit vector

$$(2) \quad \mathbf{u} = \frac{\dot{\mathbf{r}}}{|\dot{\mathbf{r}}|}$$

is called the *unit tangent vector* to C at P .

If in particular C is represented by $\mathbf{r}(s)$, where s is the arc length, it follows from (3), Sec. 5.12, that the derivative $d\mathbf{r}/ds$ is a unit vector, and (2) becomes

$$(3) \quad \mathbf{u} = \mathbf{r}' = \frac{d\mathbf{r}}{ds}.$$

Clearly the position vector of a point T on the tangent is the sum of the position vector \mathbf{r} of P and a vector in the direction of the tangent. Hence a parametric representation of the tangent is (Fig. 158)

$$(4) \quad \mathbf{q}(w) = \mathbf{r} + w\dot{\mathbf{r}}$$

where both \mathbf{r} and $\dot{\mathbf{r}}$ depend on P and the parameter w is a real variable.

Consider a given curve C , represented by a three times continuously differentiable vector function $\mathbf{r}(s)$ (cf. footnote 6) where s is the arc length.

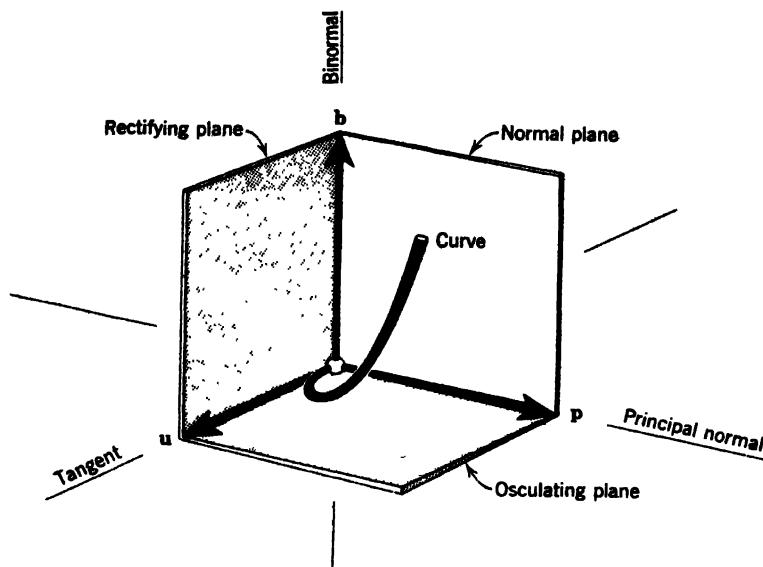


Fig. 159. Trihedron.

Then

$$(5) \quad \kappa(s) = |\mathbf{u}'(s)| = |\mathbf{r}''(s)| \quad (\kappa \geq 0)$$

is called the **curvature** of C . If $\kappa \neq 0$, the unit vector \mathbf{p} in the direction of $\mathbf{u}'(s)$ is

$$(6) \quad \mathbf{p} = \frac{\mathbf{u}'}{\kappa} \quad (\kappa > 0)$$

and is called the **unit principal normal vector** of C . From Ex. 1 in Sec. 5.10 we see that \mathbf{p} is perpendicular to \mathbf{u} . The vector

$$(7) \quad \mathbf{b} = \mathbf{u} \times \mathbf{p} \quad (\kappa > 0)$$

is called the **unit binormal vector** of C . From the definition of vector product it follows that \mathbf{u} , \mathbf{p} , \mathbf{b} constitute a right-handed triple of orthogonal unit vectors (Secs. 5.3 and 5.5). This triple is called the **trihedron** of C at the point under consideration (Fig. 159). The three straight lines through that point in the directions of \mathbf{u} , \mathbf{p} , \mathbf{b} are called the **tangent**, the **principal normal**, and the **binormal** of C . Figure 159 also shows the names of the three planes spanned by each pair of those vectors.

If the derivative \mathbf{b}' is not the zero vector, it is perpendicular to \mathbf{b} (cf. Ex. 1 in Sec. 5.10). It is also perpendicular to \mathbf{u} . In fact, by differentiating $\mathbf{b} \cdot \mathbf{u} = 0$ we have $\mathbf{b}' \cdot \mathbf{u} + \mathbf{b} \cdot \mathbf{u}' = 0$; hence $\mathbf{b}' \cdot \mathbf{u} = 0$ because $\mathbf{b} \cdot \mathbf{u}' = 0$. Consequently, \mathbf{b}' is of the form $\mathbf{b}' = \alpha \mathbf{p}$ where α is a scalar. It is customary to set $\alpha = -\tau$.

Then

$$(8) \quad \mathbf{b}' = -\tau \mathbf{p} \quad (\kappa > 0).$$

The scalar function τ is called the **torsion** of C . Scalar multiplication of both sides of (8) by p yields

$$(9) \quad \tau(s) = -p(s) \cdot b'(s).$$

Example 1 (Circular helix). In the case of the circular helix (6) in Sec. 5.11 we obtain $s = t \sqrt{a^2 + c^2}$. Hence we may represent it in the form

$$\mathbf{r}(s) = a \cos \frac{s}{w} \mathbf{i} + a \sin \frac{s}{w} \mathbf{j} + c \frac{s}{w} \mathbf{k} \quad \text{where } w = \sqrt{a^2 + c^2}.$$

It follows that

$$\mathbf{u}(s) = \mathbf{r}'(s) = -\frac{a}{w} \sin \frac{s}{w} \mathbf{i} + \frac{a}{w} \cos \frac{s}{w} \mathbf{j} + \frac{c}{w} \mathbf{k}$$

$$\mathbf{r}''(s) = -\frac{a}{w^2} \cos \frac{s}{w} \mathbf{i} - \frac{a}{w^2} \sin \frac{s}{w} \mathbf{j}$$

$$\kappa = |\mathbf{r}''| = \sqrt{\mathbf{r}'' \cdot \mathbf{r}''} = \frac{a}{w^2} = \frac{a}{a^2 + c^2}$$

$$\mathbf{p}(s) = \frac{\mathbf{r}''(s)}{\kappa(s)} = -\cos \frac{s}{w} \mathbf{i} - \sin \frac{s}{w} \mathbf{j}$$

$$\mathbf{b}(s) = \mathbf{u}(s) \times \mathbf{p}(s) = \frac{c}{w} \sin \frac{s}{w} \mathbf{i} - \frac{c}{w} \cos \frac{s}{w} \mathbf{j} + \frac{a}{w} \mathbf{k}$$

$$\mathbf{b}'(s) = \frac{c}{w^2} \cos \frac{s}{w} \mathbf{i} + \frac{c}{w^2} \sin \frac{s}{w} \mathbf{j}$$

$$\tau(s) = -\mathbf{p}(s) \cdot \mathbf{b}'(s) = \frac{c}{w^2} = \frac{c}{a^2 + c^2}.$$

Hence the circular helix has constant curvature and torsion. If $c > 0$ (right-handed helix, cf. Fig. 153), then $\tau > 0$, and if $c < 0$ (left-handed helix, cf. Fig. 154), then $\tau < 0$.

Since \mathbf{u} , \mathbf{p} , and \mathbf{b} are linearly independent vectors, we may represent any vector in space as a linear combination of these vectors. Hence if the derivatives \mathbf{u}' , \mathbf{p}' , and \mathbf{b}' exist, they may be represented in that fashion. The corresponding formulas are

$$(10) \quad \begin{aligned} (a) \quad \mathbf{u}' &= \kappa \mathbf{p} \\ (b) \quad \mathbf{p}' &= -\kappa \mathbf{u} + \tau \mathbf{b} \\ (c) \quad \mathbf{b}' &= -\tau \mathbf{p} \end{aligned}$$

They are called the **formulas of Frenet**. (10a) follows from (6), and (10c) is identical with (8). To derive (10b) we note that, by the definition of vector product,

$$\mathbf{p} = \mathbf{b} \times \mathbf{u}, \quad \mathbf{p} \times \mathbf{u} = -\mathbf{b}, \quad \mathbf{b} \times \mathbf{p} = -\mathbf{u}.$$

Differentiating the first of these formulas and using (10a) and (10c), we obtain

$$\mathbf{p}' = \mathbf{b}' \times \mathbf{u} + \mathbf{b} \times \mathbf{u}' = -\tau \mathbf{p} \times \mathbf{u} + \mathbf{b} \times \kappa \mathbf{p} = -\tau(-\mathbf{b}) + \kappa(-\mathbf{u}),$$

which proves (10b).

Problems

- Verify that \mathbf{u} , \mathbf{p} , and \mathbf{b} in Ex. 1 are orthogonal unit vectors.
- Show that the angle between \mathbf{u} in Ex. 1 and the z -axis is constant.

Find a parametric representation of the tangent of the following curves at the given point P .

- $\mathbf{r}(t) = t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k}$, $P: (1, 1, 1)$
- $\mathbf{r}(t) = \cos t\mathbf{i} + \sin t\mathbf{j} + 2t\mathbf{k}$, $P: (1, 0, 4\pi)$
- $\mathbf{r}(t) = 4 \cos t\mathbf{i} + 4 \sin t\mathbf{j}$, $P: (2\sqrt{2}, -2\sqrt{2}, 0)$

- Show that straight lines are the only curves whose unit tangent vector is constant.
- Show that if the curve C is represented by $\mathbf{r}(t)$ where t is any parameter, then the curvature is

$$(5') \quad \kappa = \frac{\sqrt{(\dot{\mathbf{r}} \cdot \ddot{\mathbf{r}})(\ddot{\mathbf{r}} \cdot \dddot{\mathbf{r}})} - (\dot{\mathbf{r}} \cdot \ddot{\mathbf{r}})^2}{(\dot{\mathbf{r}} \cdot \ddot{\mathbf{r}})^{3/2}}.$$

- Find the curvature of the ellipse $\mathbf{r}(t) = a \cos t\mathbf{i} + b \sin t\mathbf{j}$.

- Show that the curvature of a circle of radius a equals $1/a$.

- Using (5'), show that for a curve $y = y(x)$ in the xy -plane,

$$\kappa = |y''|/(1 + y'^2)^{3/2} \quad (y' = dy/dx, \text{ etc.}).$$

- Show that the curvature of a straight line is identically zero.

- Show that the torsion of a plane curve (with $\kappa > 0$) is identically zero.

- Using (7) and (9), show that

$$(9') \quad \tau = (\mathbf{u} \mathbf{p} \mathbf{p}') \quad (\kappa > 0).$$

- Using (6), show that (9') may be written

$$(9'') \quad \tau = (\mathbf{r}' \mathbf{r}'' \mathbf{r}''')/\kappa^2 \quad (\kappa > 0).$$

- Show that if the curve C is represented by $\mathbf{r}(t)$ where t is any parameter, then (9'') becomes

$$(9''') \quad \tau = \frac{(\dot{\mathbf{r}} \ddot{\mathbf{r}} \ddot{\mathbf{r}})}{(\dot{\mathbf{r}} \cdot \ddot{\mathbf{r}})(\ddot{\mathbf{r}} \cdot \ddot{\mathbf{r}}) - (\dot{\mathbf{r}} \cdot \ddot{\mathbf{r}})^2} \quad (\kappa > 0).$$

5.14 Velocity and Acceleration

Let $\mathbf{r}(t)$ be the position vector of a moving particle P in space, where t is the time. Then $\mathbf{r}(t)$ represents the path C of P . From the previous section we know that the vector

$$(1) \quad \mathbf{v} = \dot{\mathbf{r}} = \frac{d\mathbf{r}}{dt}$$

is tangent to C and, therefore, points in the instantaneous direction of motion of P . From (3) in Sec. 5.12 we see that

$$|\mathbf{v}| = \sqrt{\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}} = \frac{ds}{dt}$$

where s is the arc length, which measures the distance of P from a fixed point ($s = 0$) on C along the curve. Hence ds/dt is the speed of P . The vector \mathbf{v} is, therefore, called the **velocity vector**⁷ of the motion.

The derivative of the velocity vector is called the **acceleration vector** and will be denoted by \mathbf{a} ; thus

$$(2) \quad \mathbf{a}(t) = \dot{\mathbf{v}}(t) = \ddot{\mathbf{r}}(t).$$

Example 1 (Centripetal acceleration). The vector function

$$\mathbf{r}(t) = R \cos \omega t \mathbf{i} + R \sin \omega t \mathbf{j} \quad (\omega > 0)$$

represents a circle C of radius R with center at the origin of the xy -plane and describes a motion of a particle P in the counter-clockwise sense. The velocity vector

$$\mathbf{v} = \dot{\mathbf{r}} = -R\omega \sin \omega t \mathbf{i} + R\omega \cos \omega t \mathbf{j}$$

(cf. Fig. 160 on the next page) is tangent to C , and its magnitude, the speed

$$|\mathbf{v}| = \sqrt{\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}} = R\omega$$

is constant. The angular speed (speed divided by the distance R from the center) is equal to ω . The acceleration vector is

$$(3) \quad \mathbf{a} = \dot{\mathbf{v}} = -R\omega^2 \cos \omega t \mathbf{i} - R\omega^2 \sin \omega t \mathbf{j} = -\omega^2 \mathbf{r}.$$

We see that there is an acceleration of constant magnitude $|\mathbf{a}| = \omega^2 R$ toward the origin, the so-called **centripetal acceleration**, which results from the fact that the velocity vector is changing direction at a constant rate. The centripetal force is $m\mathbf{a}$, where m is the mass of P . The opposite vector $-m\mathbf{a}$ is called the **centrifugal force**, and the two forces are in equilibrium at each instant of the motion.

It is clear that \mathbf{a} is the rate of change of \mathbf{v} . In Ex. 1, $|\mathbf{v}| = \text{const}$, but $|\mathbf{a}| \neq 0$, which illustrates that the magnitude of \mathbf{a} is not in general the rate of change of $|\mathbf{v}|$. The reason is that in general \mathbf{a} is not tangent to the path C . In fact, by applying the chain rule of differentiation to (1), and denoting derivatives with respect to s by primes, we have

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \frac{d\mathbf{r}}{ds} \frac{ds}{dt} = \mathbf{r}' \frac{ds}{dt}$$

and by differentiating this again,

$$(4) \quad \mathbf{a} = \frac{d\mathbf{v}}{dt} = \frac{d}{dt} \left(\mathbf{r}' \frac{ds}{dt} \right) = \mathbf{r}'' \left(\frac{ds}{dt} \right)^2 + \mathbf{r}' \frac{d^2 s}{dt^2}.$$

Since \mathbf{r}' is the unit tangent vector \mathbf{u} to C (Sec. 5.13) and its derivative $\mathbf{u}' = \mathbf{r}''$ is perpendicular to \mathbf{u} (Sec. 5.13), the formula (4) is a decomposition of the acceleration vector into its normal component $\mathbf{r}'' \mathbf{s}^2$ and its tangential component $\mathbf{r}' \mathbf{s}'$. From this we see that if, and only if, the normal component is zero, $|\mathbf{a}|$ equals the time rate of change of $|\mathbf{v}| = \mathbf{s}$ (except for the sign), because then $|\mathbf{a}| = |\mathbf{r}'| |\mathbf{s}'| = |\mathbf{s}'|$.

⁷ When no confusion is likely to arise, the word "velocity" is often used to denote the speed, the magnitude of \mathbf{v} .

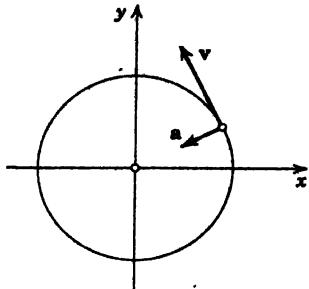


Fig. 160. Centripetal acceleration.

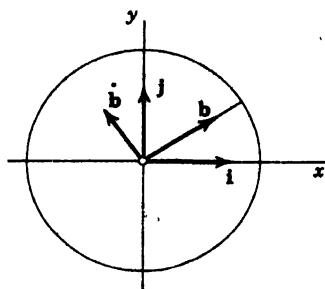


Fig. 161. Motion in Example 2.

Example 2 (Coriolis⁸ acceleration). A particle P moves on a straight line from the center of a disk towards the edge, the position vector being

$$(5) \quad \mathbf{r}(t) = t\mathbf{b}$$

where \mathbf{b} is a unit vector, rotating together with the disk with constant angular speed ω in the counterclockwise sense (Fig. 161). Find the acceleration \mathbf{a} of P .

Because of the rotation, \mathbf{b} is of the form

$$(6) \quad \mathbf{b}(t) = \cos \omega t \mathbf{i} + \sin \omega t \mathbf{j}.$$

Differentiating (5), we obtain the velocity

$$(7) \quad \mathbf{v} = \dot{\mathbf{r}} = \mathbf{b} + t\dot{\mathbf{b}}.$$

Obviously \mathbf{b} is the velocity of P relative to the disk, and $t\dot{\mathbf{b}}$ is the additional velocity due to the rotation. Differentiating once more, we obtain the acceleration

$$(8) \quad \mathbf{a} = \ddot{\mathbf{r}} = 2\dot{\mathbf{b}} + t\ddot{\mathbf{b}}.$$

In the last term of (8) we have $\ddot{\mathbf{b}} = -\omega^2\mathbf{b}$, as follows by differentiating (6). Hence this acceleration $t\ddot{\mathbf{b}}$ is directed toward the center of the disk, and from Ex. 1 we see that this is the centripetal acceleration due to the rotation. In fact, the distance of P from the center is equal to t which, therefore, plays the role of R in Ex. 1.

The most interesting and probably unexpected term in (8) is $2\dot{\mathbf{b}}$, the so-called *Coriolis acceleration*, which results from the interaction of the rotation of the disk and the motion of P on the disk. It has the direction of $\dot{\mathbf{b}}$, that is, it is tangential to the edge of the disk and, referred to the fixed xy -coordinate system, it points in the direction of the rotation. If P is a person of mass m_0 walking on the disk according to (5), he will feel a force $-2m_0\dot{\mathbf{b}}$ in the opposite direction, that is, against the sense of rotation.

Example 3 (Superposition of two rotations). Find the acceleration of a particle P moving on a "meridian" M of a rotating sphere with constant speed relative to the sphere.

The motion of P on M can be described analytically in the form

$$(9) \quad \mathbf{r}(t) = R \cos \gamma t \mathbf{b} + R \sin \gamma t \mathbf{k}$$

where R is the radius of the sphere, $\gamma (>0)$ the angular speed of P on M , \mathbf{b} a horizontal unit vector in the plane of M (Fig. 162), and \mathbf{k} the unit vector in the positive z -direction. Since \mathbf{b} rotates together with the sphere, it is of the form

$$(10) \quad \mathbf{b} = \cos \omega t \mathbf{i} + \sin \omega t \mathbf{j}$$

⁸ GUSTAVE GASPARD CORIOLIS (1792–1843), French physicist.

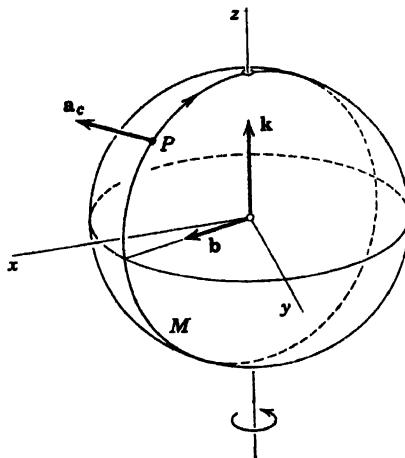


Fig. 162. Superposition of two rotations.

where $\omega (>0)$ is the angular speed of the sphere and \mathbf{i} and \mathbf{j} are the unit vectors in the positive x and y directions, which are fixed in space. By differentiating (9) we obtain the velocity

$$(11) \quad \mathbf{v} = \dot{\mathbf{r}} = R \cos \gamma t \mathbf{b} - \gamma R \sin \gamma t \mathbf{b} + \gamma R \cos \gamma t \mathbf{k}.$$

By differentiating this again we obtain the acceleration

$$(12) \quad \mathbf{a} = \ddot{\mathbf{v}} = R \cos \gamma t \ddot{\mathbf{b}} - 2\gamma R \sin \gamma t \dot{\mathbf{b}} - \gamma^2 R \cos \gamma t \mathbf{b} - \gamma^2 R \sin \gamma t \mathbf{k},$$

where, by (10),

$$\begin{aligned} \dot{\mathbf{b}} &= -\omega \sin \omega t \mathbf{i} + \omega \cos \omega t \mathbf{j}, \\ \ddot{\mathbf{b}} &= -\omega^2 \cos \omega t \mathbf{i} - \omega^2 \sin \omega t \mathbf{j} = -\omega^2 \mathbf{b}. \end{aligned}$$

From (9) we see that the sum of the last two terms in (12) is equal to $-\gamma^2 \mathbf{r}$, and (12) becomes

$$(13) \quad \mathbf{a} = -\omega^2 R \cos \gamma t \mathbf{b} - 2\gamma R \sin \gamma t \dot{\mathbf{b}} - \gamma^2 \mathbf{r}.$$

The first term on the right is the centripetal acceleration caused by the rotation of the sphere, and the last term is the centripetal acceleration resulting from the rotation of P on M . The second term is the *Coriolis acceleration*

$$(14) \quad \mathbf{a}_c = -2\gamma R \sin \gamma t \dot{\mathbf{b}}.$$

On the "Northern hemisphere," $\sin \gamma t > 0$ [cf. (9)] and because of the minus sign, \mathbf{a}_c is directed opposite to $\dot{\mathbf{b}}$, that is, tangential to the surface of the sphere, perpendicular to M , and opposite to the rotation of the sphere. Its magnitude $2\gamma R |\sin \gamma t| \omega$ is maximum at the "North pole" and zero at the equator. If P is a fly of mass m_0 walking according to (9), it will feel a force $-m_0 \mathbf{a}_c$, opposite to $m_0 \dot{\mathbf{b}}$; this is quite similar to the force felt in Ex. 2. This force tends to let the fly deviate from the path M to the right. On the "Southern hemisphere," $\sin \gamma t < 0$ and that force acts in the opposite direction, tending to let the fly deviate from M to the left. This effect can be observed in connection with missiles and other projectiles. The flow of the air towards an area of low pressure also shows these deviations.

Problems

1. Obtain (3) by differentiating (7), Sec. 5.6.
2. Find the Coriolis acceleration in Ex. 2 with (5) replaced by $\mathbf{r} = t^2 \mathbf{b}$.

Let $\mathbf{r}(t)$ be the position vector of a moving particle where $t (\geq 0)$ is the time. Describe the geometric shape of the path and find the velocity vector, the speed, and the acceleration vector.

- | | |
|---|--|
| 3. $\mathbf{r} = t \mathbf{i}$ | 4. $\mathbf{r} = t^2 \mathbf{k}$ |
| 5. $\mathbf{r} = (2t - t^2) \mathbf{i}$ | 6. $\mathbf{r} = t^3 \mathbf{i}$ |
| 7. $\mathbf{r} = \sin t \mathbf{j}$ | 8. $\mathbf{r} = 3 \cos 2t \mathbf{i} + 3 \sin 2t \mathbf{j}$ |
| 9. $\mathbf{r} = \cos t^2 \mathbf{i} + \sin t^2 \mathbf{j}$ | 10. $\mathbf{r} = 3 \cos 2t \mathbf{i} + 2 \sin 3t \mathbf{j}$ |

5.15 Chain Rule and Mean Value Theorem for Functions of Several Variables

We shall now consider some facts about functions of several variables which will be needed in the following sections. For the sake of simplicity we shall formulate everything for functions of two variables; the generalization to functions of three or more variables will be obvious. Familiarity of the student with the content of Sec. 0.2 on partial derivatives will be assumed.

A function $f(x, y)$ is said to be **continuous at a point** (x_0, y_0) if f is defined in a *neighborhood*⁹ of that point and if for any positive number ϵ (no matter how small, but not zero) we can find a positive number δ such that

$$|f(x, y) - f(x_0, y_0)| < \epsilon$$

for all (x, y) for which

$$(x - x_0)^2 + (y - y_0)^2 < \delta^2.$$

Geometrically speaking, continuity of $f(x, y)$ at (x_0, y_0) means that for each interval of length 2ϵ with midpoint $f(x_0, y_0)$ we can find a circular disk with nonzero radius δ and center (x_0, y_0) in the xy -plane such that for every point (x, y) in the disk the corresponding function value $f(x, y)$ lies in that interval (Fig. 163).

From elementary calculus we know that if w is a differentiable function of x , and x is a differentiable function of t , then

$$\frac{dw}{dt} = \frac{dw}{dx} \frac{dx}{dt}.$$

⁹ That is, in some circular disk $(x - x_0)^2 + (y - y_0)^2 < r^2$, ($r > 0$).

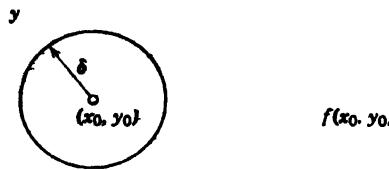


Fig. 163. Continuity of a function of two variables.

This so-called chain rule of differentiation can be generalized as follows.

Theorem 1 (Chain rule). Let $w = f(x, y)$ be continuous and have continuous first partial derivatives at every point of a domain¹⁰ D in the xy -plane. Let $x = x(t)$ and $y = y(t)$ be differentiable functions of a variable t in some interval T such that, for each t in T , the point $[x(t), y(t)]$ lies in D . Then $w = f[x(t), y(t)]$ is a differentiable function for all t in T , and

$$(1) \quad \frac{dw}{dt} = \frac{\partial w}{\partial x} \frac{dx}{dt} + \frac{\partial w}{\partial y} \frac{dy}{dt}.$$

Proof. We choose a t in T and Δt so small that $t + \Delta t$ is also in T , and set

$$(2) \quad \Delta x = x(t + \Delta t) - x(t), \quad \Delta y = y(t + \Delta t) - y(t),$$

and furthermore

$$\Delta w = f(x + \Delta x, y + \Delta y) - f(x, y).$$

By adding and subtracting a term this may be written

$$\Delta w = [f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)] + [f(x, y + \Delta y) - f(x, y)].$$

If we apply the mean value theorem for a function of a single variable (cf. Ref. [A13]) to each of the two expressions in brackets, we obtain

$$(3) \quad \Delta w = \Delta x \frac{\partial f}{\partial x}_{x_1, y_1 + \Delta y} + \Delta y \frac{\partial f}{\partial y}$$

where x_1 lies between x and $x + \Delta x$, and y_1 lies between y and $y + \Delta y$. Dividing (3) by Δt on both sides, letting Δt approach 0, and noting that $\partial f / \partial x$ and $\partial f / \partial y$ are assumed to be continuous, we obtain (1).

This theorem may now immediately be extended as follows.

Theorem 2. Let $w = f(x, y)$ be continuous and have continuous first partial derivatives in a domain D of the xy -plane. Let $x = x(u, v)$ and $y = y(u, v)$ be functions which have first partial derivatives in a domain B of the uv -plane, which is such that, for any point (u, v) in B , the corresponding point

¹⁰ A domain D is an open connected point set, where "connected" means that any two points of D can be joined by a broken line of finitely many linear segments all of whose points belong to D , and "open" means that every point of D has a neighborhood all of whose points belong to D . For example, the interior of a rectangle or a circle is a domain.

$[x(u, v), y(u, v)]$ lies in D . Then the function $w = f(x(u, v), y(u, v))$ is defined and has first partial derivatives in B , and

$$(4) \quad \begin{aligned}\frac{\partial w}{\partial u} &= \frac{\partial w}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial u}, \\ \frac{\partial w}{\partial v} &= \frac{\partial w}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial v}.\end{aligned}$$

The proof follows immediately from Theorem 1 by keeping one of the two variables u and v constant.

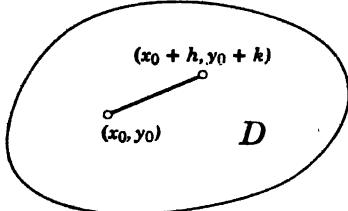


Fig. 164. Mean value theorem.

From elementary calculus we know that if a function $f(x)$ is differentiable, then

$$f(x_0 + h) - f(x_0) = h \frac{df}{dx},$$

the derivative being evaluated at a suitable point between x_0 and $x_0 + h$ (cf. Ref. [A13]). This so-called mean-value theorem

of differential calculus can be extended to functions of two variables as follows.

Theorem 3 (Mean-value theorem). Let $f(x, y)$ be continuous and have continuous first partial derivatives in a domain D . Furthermore, let (x_0, y_0) and $(x_0 + h, y_0 + k)$ be points in D such that the line segment joining these points lies in D (Fig. 164). Then

$$(5) \quad f(x_0 + h, y_0 + k) - f(x_0, y_0) = h \frac{\partial f}{\partial x} + k \frac{\partial f}{\partial y},$$

the partial derivatives being evaluated at a suitable point of that segment.

Proof. Let

$$x = x_0 + th, \quad y = y_0 + tk \quad (0 \leq t \leq 1)$$

and, furthermore,

$$F(t) = f(x_0 + th, y_0 + tk).$$

Then

$$f(x_0 + h, y_0 + k) = F(1), \quad f(x_0, y_0) = F(0)$$

and, by the mean-value theorem for a function of a single variable,

$$(6) \quad f(x_0 + h, y_0 + k) - f(x_0, y_0) = F(1) - F(0) = F'(t_1)$$

where t_1 is a value between 0 and 1. Since $dx/dt = h$ and $dy/dt = k$, we obtain from Theorem 1

$$(7) \quad F'(t_1) = \frac{\partial f}{\partial x} h + \frac{\partial f}{\partial y} k.$$



the derivatives on the right being evaluated at the point $(x_0 + t_1 h, y_0 + t_1 k)$, which lies on the line segment with end points (x_0, y_0) and $(x_0 + h, y_0 + k)$. By inserting (7) into (6) we obtain (5), and the proof is complete.

For a function $f(x, y, z)$ of three variables, satisfying conditions analogous to those in Theorem 3, the consideration is quite similar and leads to the formula

$$(8) \quad f(x_0 + h, y_0 + k, z_0 + l) - f(x_0, y_0, z_0) = h \frac{\partial f}{\partial x} + k \frac{\partial f}{\partial y} + l \frac{\partial f}{\partial z},$$

the partial derivatives being evaluated at a suitable point of the segment with end points (x_0, y_0, z_0) and $(x_0 + h, y_0 + k, z_0 + l)$.

Problems

Using (1), find dw/dt where

$$\begin{array}{ll} 1. w = x + y, & x = t^2, \quad y = \ln t \\ 3. w = x/y, & x = g(t), \quad y = h(t) \end{array} \quad \begin{array}{ll} 2. w = \sqrt{x^2 + y^2}, & x = e^{2t}, \quad y = e^{-2t} \\ 4. w = x^y, & x = \cos t, \quad y = \sin t \end{array}$$

5. Let $w = f(x, y, z)$ where x, y, z are functions of t . Show that under conditions similar to those in Theorem 1,

$$(9) \quad \frac{dw}{dt} = \frac{\partial w}{\partial x} \frac{dx}{dt} + \frac{\partial w}{\partial y} \frac{dy}{dt} + \frac{\partial w}{\partial z} \frac{dz}{dt}.$$

Using (9), find dw/dt where

$$\begin{array}{ll} 6. w = x^2 + y^2 + z^2, & x = e^{2t} \cos 2t, \quad y = e^{2t} \sin 2t, \quad z = e^{4t} \\ 7. w = (x^2 + y^2 + z^2)^{-1/2}, & x = \cos t, \quad y = \sin t, \quad z = t \end{array}$$

8. Prove Theorem 2.

Find $\partial w/\partial u$ and $\partial w/\partial v$ where

$$\begin{array}{l} 9. w = \ln(x^2 + y^2), \quad x = e^u \cos v, \quad y = e^u \sin v \\ 10. w = x^2 - y^2, \quad x = u^2 - v^2, \quad y = 2uv \\ 11. w = xy, \quad x = e^u \cos v, \quad y = e^u \sin v \end{array}$$

12. Derive (8).

13. Let $w = f(x, y)$ and $x = r \cos \theta, y = r \sin \theta$. Show that

$$\left(\frac{\partial w}{\partial r} \right)^2 + \frac{1}{r^2} \left(\frac{\partial w}{\partial \theta} \right)^2 = \left(\frac{\partial w}{\partial x} \right)^2 + \left(\frac{\partial w}{\partial y} \right)^2.$$

14. Let $w = f(v, z)$ and $v = x + ct, z = x - ct$, where c is a constant. Show that, granted sufficient differentiability,

$$c^2 w_{xx} - w_{tt} = 4c^2 w_{vz}.$$

15. Let $w = f(x, y)$ and $x = r \cos \theta, y = r \sin \theta$. Show that

$$w_{xx} + w_{yy} = w_{rr} + \frac{1}{r} w_r + \frac{1}{r^2} w_{\theta\theta}.$$

5.16 Directional Derivative. Gradient of a Scalar Field

We consider a scalar field in space (Sec. 5.9) given by a scalar function $f(P) = f(x, y, z)$. We know that the first partial derivatives of f are the rates of change of f in the directions of the coordinate axes. It seems unnatural to restrict attention to these three directions, and we may ask for the rate of change of f in any direction. This simple idea leads to the notion of a directional derivative.

To define that derivative we choose a point P in space and a direction at P , given by a unit vector \mathbf{b} . Let C be the ray from P in the direction of \mathbf{b} , and let Q be a point on C , whose distance from P is s (Fig. 165). Then if the limit

$$(1) \quad \frac{\partial f}{\partial s} = \lim_{s \rightarrow 0} \frac{f(Q) - f(P)}{s} \quad (s = \text{distance between } P \text{ and } Q)$$

exists, it is called the **directional derivative of f at P in the direction of \mathbf{b}** . Obviously, $\partial f / \partial s$ is the rate of change of f at P in the direction of \mathbf{b} .

In this way there are now infinitely many directional derivatives of f at P , each corresponding to a certain direction. But, a Cartesian coordinate system being given, we may represent any such derivative in terms of the first partial derivatives of f at P as follows. If P has the position vector \mathbf{a} , then the ray C can be represented in the form

$$(2) \quad \mathbf{r}(s) = x(s)\mathbf{i} + y(s)\mathbf{j} + z(s)\mathbf{k} = \mathbf{a} + s\mathbf{b} \quad (s \geq 0),$$

and $\partial f / \partial s$ is the derivative of the function $f[x(s), y(s), z(s)]$ with respect to the arc length s of C . Hence, assuming that f has continuous first partial derivatives and applying the chain rule (Theorem 1 in the last section), we obtain

$$(3) \quad \frac{\partial f}{\partial s} = \frac{\partial f}{\partial x} x' + \frac{\partial f}{\partial y} y' + \frac{\partial f}{\partial z} z'$$

where primes denote derivatives with respect to s . Now from (2),

$$\mathbf{r}' = x'\mathbf{i} + y'\mathbf{j} + z'\mathbf{k} = \mathbf{b}.$$

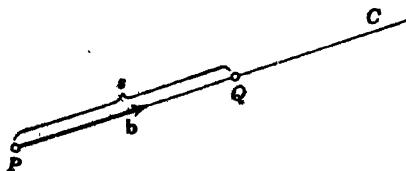


Fig. 165. Directional derivative.

This suggests that we introduce the vector

$$(4) \quad \text{grad } f = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} + \frac{\partial f}{\partial z} \mathbf{k}$$

and write (3) in the form of a scalar product:

$$(5) \quad \frac{\partial f}{\partial s} = \mathbf{b} \cdot \text{grad } f \quad (|\mathbf{b}| = 1).$$

The vector $\text{grad } f$ is called the **gradient** of the scalar function f .

By introducing the *differential operator*

$$\nabla = \frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial y} \mathbf{j} + \frac{\partial}{\partial z} \mathbf{k}$$

(read **nabla** or "del") we may write

$$\text{grad } f = \nabla f = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} + \frac{\partial f}{\partial z} \mathbf{k}.$$

The notation ∇f for the gradient is frequently used in the engineering literature.

We note that if in particular \mathbf{b} has the direction of the positive x -axis, then $\mathbf{b} = \mathbf{i}$, and

$$\frac{\partial f}{\partial s} = \mathbf{b} \cdot \text{grad } f = \frac{\partial f}{\partial x} \mathbf{i} \cdot \mathbf{i} = \frac{\partial f}{\partial x}.$$

Similarly, the directional derivative in the positive y -direction is $\partial f / \partial y$, etc.

We want to show that $\text{grad } f$ is a vector, that is, has a magnitude and direction which is independent of the choice of coordinates in space.

This is, of course, not obvious; the definition (4) merely shows that $\text{grad } f$ looks like a vector, but since (4) involves partial derivatives, which depend on the choice of the coordinates, we do not know yet whether the corresponding expression

$$\frac{\partial f}{\partial x^*} \mathbf{i}^* + \frac{\partial f}{\partial y^*} \mathbf{j}^* + \frac{\partial f}{\partial z^*} \mathbf{k}^*$$

with respect to other Cartesian coordinates x^*, y^*, z^* (and corresponding unit vectors $\mathbf{i}^*, \mathbf{j}^*, \mathbf{k}^*$) will have the same magnitude and direction as (4). For comparison, the expression

$$\frac{\partial f}{\partial x} \mathbf{i} + 2 \frac{\partial f}{\partial y} \mathbf{j} + \frac{\partial f}{\partial z} \mathbf{k}$$

looks like a vector too, but the corresponding expression in another coordinate system will generally have another magnitude and direction, and the expression, therefore, is not a vector.

To show that $\text{grad } f$ is a vector, we may reason as follows. By definition of a scalar function, the value of f at a point P depends on P but is independent of the coordinates, and s , the arc length of that ray C , has the same property. Hence $\partial f / \partial s$ is independent of the particular choice of coordinates. From (5) we obtain

$$\frac{\partial f}{\partial s} = |\mathbf{b}| |\text{grad } f| \cos \gamma = |\text{grad } f| \cos \gamma$$

where γ is the angle between \mathbf{b} and $\text{grad } f$. We see that $\partial f / \partial s$ is maximum when $\cos \gamma = 1$, $\gamma = 0$, and then $\partial f / \partial s = |\text{grad } f|$. This shows that the magnitude and direction of $\text{grad } f$ are independent of the coordinates, and we have the following result.

Theorem 1. *Let $f(P) = f(x, y, z)$ be a scalar function having continuous first partial derivatives. Then $\text{grad } f$ exists and is a vector. If at a point P the gradient of f is not the zero vector, it has the direction of maximum increase of f at P .*

Another important geometrical characterization of the gradient can be obtained as follows. Consider a scalar function $f(x, y, z)$ and suppose that for each constant c the equation

$$(6) \quad f(x, y, z) = c = \text{const}$$

represents a surface S in space. Then, by letting c assume all values, we obtain a family of surfaces, which are called the **level surfaces** of the function f . We shall assume that f is such that through each point in space there passes one, and only one, level surface of f . We remember that a curve C in space may be represented in the form (cf. Sec. 5.11)

$$(7) \quad \mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}.$$

If we now require that C lies on S , then the functions $x(t)$, $y(t)$, and $z(t)$ in (7) must be such that

$$f[x(t), y(t), z(t)] = c;$$

cf. (6). By differentiating this with respect to t and using the chain rule (Sec. 5.15) we obtain

$$(8) \quad \frac{\partial f}{\partial x} \dot{x} + \frac{\partial f}{\partial y} \dot{y} + \frac{\partial f}{\partial z} \dot{z} = (\text{grad } f) \cdot \dot{\mathbf{r}} = 0,$$

where the vector

$$\dot{\mathbf{r}} = \dot{x}\mathbf{i} + \dot{y}\mathbf{j} + \dot{z}\mathbf{k}$$

is tangent to C (cf. Sec. 5.13). If we consider curves on S passing through a point P of S in various directions, their tangents at P will, in general, lie in the same plane which touches S at P . This plane is called the **tangent plane** of S at P . The straight line through P and perpendicular to the tangent plane is called the **normal** of S at P (Fig. 166). From (8) and Theorem 1, Sec. 5.4, we thus obtain the following result.

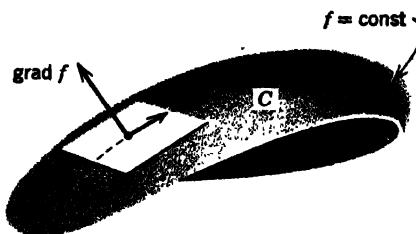


Fig. 166. Level surface and gradient.

Theorem 2. If f is a scalar function such that through a point P in space there passes precisely one level surface S of f , and $\text{grad } f \neq 0$ at P , then $\text{grad } f$ is perpendicular to S at P , that is, has the direction of the normal of S at P .

Example 1. Find the directional derivative $\partial f / \partial s$ of $f(x, y, z) = 2x^2 + 3y^2 + z^2$ at the point $P: (2, 1, 3)$ in the direction of the vector $\mathbf{a} = \mathbf{i} - 2\mathbf{k}$. We obtain

$$\text{grad } f = 4x\mathbf{i} + 6y\mathbf{j} + 2z\mathbf{k}, \quad \text{and at } P, \quad \text{grad } f = 8\mathbf{i} + 6\mathbf{j} + 6\mathbf{k}.$$

Since $|\mathbf{a}| = \sqrt{5}$, the unit vector in the direction of \mathbf{a} is

$$\mathbf{b} = \frac{\mathbf{a}}{|\mathbf{a}|} = \frac{1}{\sqrt{5}}\mathbf{i} - \frac{2}{\sqrt{5}}\mathbf{k}.$$

Therefore,

$$\frac{\partial f}{\partial s} = (8\mathbf{i} + 6\mathbf{j} + 6\mathbf{k}) \cdot \left(\frac{1}{\sqrt{5}}\mathbf{i} - \frac{2}{\sqrt{5}}\mathbf{k} \right) = -\frac{4}{\sqrt{5}}.$$

The minus sign indicates that f decreases in the direction under consideration.

Example 2 (Normal to a plane curve). The level curves $f = \text{const}$ of the function $f(x, y) = \ln(x^2 + y^2)$ are concentric circles about the origin. The gradient

$$\text{grad } f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} = \frac{2x}{x^2 + y^2}\mathbf{i} + \frac{2y}{x^2 + y^2}\mathbf{j}$$

has the direction of the normals to the circles, and its direction corresponds to that of the maximum increase of f . For example, at the point $P: (2, 1)$, we have (cf. Fig. 167 on the next page).

$$\text{grad } f = 0.8\mathbf{i} + 0.4\mathbf{j}$$

Example 3 (Normal to a surface). Find a unit normal vector \mathbf{n} of the cone of revolution $z^2 = 4(x^2 + y^2)$ at the point $P: (1, 0, 2)$. We may regard the cone as the level surface $f = 0$ of the function $f(x, y, z) = 4(x^2 + y^2) - z^2$. Then

$$\text{grad } f = 8x\mathbf{i} + 8y\mathbf{j} - 2z\mathbf{k} \quad \text{and at } P, \quad \text{grad } f = 8\mathbf{i} - 4\mathbf{k}.$$

Hence, by Theorem 2,

$$\mathbf{n} = \frac{\text{grad } f}{|\text{grad } f|} = \frac{2}{\sqrt{5}}\mathbf{i} - \frac{1}{\sqrt{5}}\mathbf{k}$$

is a unit normal vector of the cone at P , and $-\mathbf{n}$ is the other one (cf. Fig. 168 on the next page).

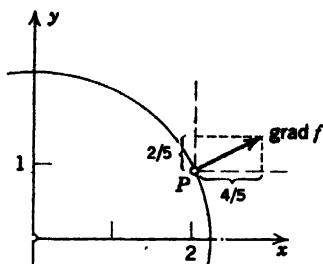
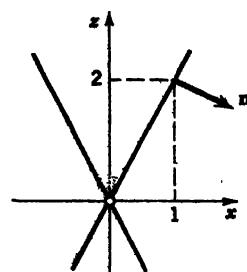


Fig. 167. Normal to a circle.

Fig. 168. Intersection of the xz -plane and the cone in Example 3.

Some of the vector fields occurring in physics are given by vector functions which can be obtained as the gradients of suitable scalar functions. Such a scalar function is then called the *potential function* or *potential* of the corresponding vector field. The use of potentials simplifies the investigation of those vector fields considerably. To obtain a first impression of this approach to vector fields, let us consider an important example.

Example 4 (Gravitational field. Laplace's equation). In Ex. 4, Sec. 5.9, we have seen that, according to Newton's law of gravitation, the force of attraction between two particles is

$$(9) \quad \mathbf{p} = -c \frac{\mathbf{r}}{r^3} = -c \left(\frac{x - x_0}{r^3} \mathbf{i} + \frac{y - y_0}{r^3} \mathbf{j} + \frac{z - z_0}{r^3} \mathbf{k} \right)$$

where

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$$

is the distance between the two particles and c is a constant. Observing that

$$(10a) \quad \frac{\partial}{\partial x} \left(\frac{1}{r} \right) = - \frac{2(x - x_0)}{2[(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2]^{3/2}} = - \frac{x - x_0}{r^3}$$

and similarly

$$(10b) \quad \frac{\partial}{\partial y} \left(\frac{1}{r} \right) = - \frac{y - y_0}{r^3}, \quad \frac{\partial}{\partial z} \left(\frac{1}{r} \right) = - \frac{z - z_0}{r^3},$$

we see that \mathbf{p} is the gradient of the scalar function

$$f(x, y, z) = \frac{c}{r} \quad (r > 0);$$

that is, f is the potential of that gravitational field.

By differentiating (10) we find

$$\begin{aligned} \frac{\partial^2}{\partial x^2} \left(\frac{1}{r} \right) &= - \frac{1}{r^3} + \frac{3(x - x_0)^2}{r^5}, & \frac{\partial^2}{\partial y^2} \left(\frac{1}{r} \right) &= - \frac{1}{r^3} + \frac{3(y - y_0)^2}{r^5}, \\ \frac{\partial^2}{\partial z^2} \left(\frac{1}{r} \right) &= - \frac{1}{r^3} + \frac{3(z - z_0)^2}{r^5}. \end{aligned}$$

Since the sum of the three expressions on the right is zero, we see that the potential $f = c/r$ satisfies the equation

$$(11) \quad \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0.$$

This important partial differential equation is called **Laplace's equation**; it will be considered in detail in Chaps. 9 and 16. The expression on the left is called the **Laplacian of f** and is denoted by $\nabla^2 f$ or Δf . The differential operator

$$\nabla^2 = \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

(read "nabla squared" or "delta") is called the **Laplace operator**. Using this operator, we may write (11) in the form

$$\nabla^2 f = 0.$$

It can be shown that the field of force produced by any distribution of masses is given by a vector function which is the gradient of a scalar function f , and f satisfies (11) in any region of space which is free of matter.

There are other laws in physics which are of the same form as Newton's law of gravitation. For example, in electrostatics the force of attraction (or repulsion) between two particles of opposite (or like) charges Q_1 and Q_2 is

$$\mathbf{p} = k \frac{\mathbf{r}}{r^3} \quad (\text{Coulomb's law}),$$

where $k = Q_1 Q_2 / 4\pi\epsilon$, and ϵ is the dielectric constant. Hence \mathbf{p} is the gradient of the potential $f = -k/r$, and f satisfies (11).

If the vector function defining a vector field is the gradient of a scalar function, the field is said to be **conservative**, because, as we shall see in Sec. 6.12, in such a field the work done in displacing a particle from a point P_1 to a point P_2 in the field depends only on P_1 and P_2 but not on the path along which the particle is displaced from P_1 to P_2 . We shall see that not every field is conservative.

Problems

Find the gradient ∇f , where

- | | | |
|------------------------------------|-------------------------|---------------------------------|
| 1. $f = 2x - y$ | 2. $f = e^x \cos y$ | 3. $f = \sin x \cosh y$ |
| 4. $f = x^2 - y^2$ | 5. $f = \ln(x^2 + y^2)$ | 6. $f = \arctan(y/x)$ |
| 7. $f = xyz$ | 8. $f = yz + zx + xy$ | 9. $f = x^2 + y^2 + z^2$ |
| 10. $f = (x^2 + y^2 + z^2)^{-1/2}$ | 11. $f = e^{xyz}$ | 12. $f = \cos(x^2 + y^2 + z^2)$ |

Graph some level curves $f = \text{const}$. Find ∇f and indicate ∇f by arrows at some points of the level curves.

- | | | |
|-------------------|----------------------|-----------------------|
| 13. $f = 2x - 3y$ | 14. $f = y/x$ | 15. $f = xy$ |
| 16. $f = x^2y^2$ | 17. $f = 9x^2 + y^2$ | 18. $f = 4x^2 + 9y^2$ |

Find and graph a normal vector to the given plane curve at a point $P: (x, y)$.

- | | |
|---------------------------------|-------------------------------------|
| 19. $y = x$, $P: (3, 3)$ | 20. $y = x^2$, $P: (-2, 4)$ |
| 21. $y = 3x - 4$, $P: (1, -1)$ | 22. $x^2 + y^2 = 25$, $P: (3, 4)$ |
| 23. $y^2 = x^3$, $P: (4, -8)$ | 24. $3x^2 - 2y^2 = 1$, $P: (1, 1)$ |

Find a normal vector to the given surface at a point $P: (x, y, z)$.

- | | |
|--|---|
| 25. $x + y + z = 0$, $P: (1, 1, -2)$ | 26. $2x + 3y - z = 1$, $P: (1, 2, 7)$ |
| 27. $z = x^2 + y^2$, $P: (2, 2, 8)$ | 28. $x^2 + y^2 + z^2 = 4$, $P: (2, -2, 0)$ |
| 29. $x^2 + 3y^2 + 2z^2 = 6$, $P: (2, 0, 1)$ | 30. $z = xy$, $P: (3, -1, -3)$ |

Find a scalar function f such that $\mathbf{v} = \nabla f$.

31. $\mathbf{v} = \mathbf{i} - \mathbf{j} + \mathbf{k}$

32. $\mathbf{v} = x\mathbf{i} + y\mathbf{j} + \mathbf{k}$

33. $\mathbf{v} = x\mathbf{i} + 2y\mathbf{j} + z\mathbf{k}$

34. $\mathbf{v} = yz\mathbf{i} + zx\mathbf{j} + xy\mathbf{k}$

35. $\mathbf{v} = (xi + yj)/(x^2 + y^2)$

36. $\mathbf{v} = e^x \sin y \mathbf{i} + e^x \cos y \mathbf{j}$

37. Find the directional derivative of $f = x^2 + y^2$ at the point $P: (2, 2)$ in the directions of the vectors $\mathbf{i}, \mathbf{i} + \mathbf{j}, \mathbf{j}, -\mathbf{i} + \mathbf{j}, -\mathbf{i}, -\mathbf{i} - \mathbf{j}, -\mathbf{j}$, and $\mathbf{i} - \mathbf{j}$.

Find the directional derivative of f at P in the direction of \mathbf{a} , where

38. $f = x^2 - y^2, P: (2, 3), \mathbf{a} = \mathbf{i} + \mathbf{j}$

39. $f = 2x + 3y, P: (0, 2), \mathbf{a} = 3\mathbf{j}$

40. $f = x^2 + y^2, P: (1, 2), \mathbf{a} = \mathbf{i} - \mathbf{j}$

41. $f = y/x, P: (3, -1), \mathbf{a} = 9\mathbf{i} - 3\mathbf{j}$

42. $f = x + 2y - z, P: (1, 4, 0), \mathbf{a} = \mathbf{j} - \mathbf{k}$

43. $f = x^2 + y^2 + z^2, P: (2, 0, 3), \mathbf{a} = 2\mathbf{i} - \mathbf{j}$

44. $f = xyz, P: (3, 2, -1), \mathbf{a} = 3\mathbf{i} + 2\mathbf{j} - 2\mathbf{k}$

45. $f = 1/(x^2 + y^2 + z^2), P: (1, 2, -3), \mathbf{a} = 2\mathbf{i} - 3\mathbf{j} + \mathbf{k}$

46. What are the directional derivatives of a function $f(x, y, z)$ in the x , y , and z directions? In the negative x , y , and z directions?

47. Let r be the distance from a fixed point P to a point $Q: (x, y, z)$. What are the level surfaces of r ? Show that ∇r is a unit vector in the direction from P to Q .

48. Show that the functions in Prob. 1-6 are solutions of Laplace's equation.

49. Show that $\nabla(fg) = f\nabla g + g\nabla f$.

50. Show that $\nabla^2(fg) = g\nabla^2f + 2\nabla f \cdot \nabla g + f\nabla^2g$.

5.17 Transformation of Coordinate Systems and Vector Components

We shall now characterize the transformations which carry Cartesian coordinate systems into Cartesian coordinate systems and investigate the change of vector components under such a transformation. This problem is basic for theoretical as well as practical reasons.

Let x, y, z and x^*, y^*, z^* be any two systems of Cartesian coordinates. Let

(1) (a) $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ and (b) $\mathbf{v} = v_1^*\mathbf{i}^* + v_2^*\mathbf{j}^* + v_3^*\mathbf{k}^*$

be the representations of a given vector \mathbf{v} in these two coordinate systems; here $\mathbf{i}, \mathbf{j}, \mathbf{k}$ and $\mathbf{i}^*, \mathbf{j}^*, \mathbf{k}^*$ are unit vectors in the positive x, y, z and x^*, y^*, z^* directions, respectively. We want to express the components v_1^*, v_2^*, v_3^* in terms of the components v_1, v_2, v_3 , and conversely.

From (1a) it follows that

(2) $\mathbf{i}^* \cdot \mathbf{v} = v_1\mathbf{i}^* \cdot \mathbf{i} + v_2\mathbf{i}^* \cdot \mathbf{j} + v_3\mathbf{i}^* \cdot \mathbf{k}$.

Similarly, by taking the scalar product of (1b) and \mathbf{i}^* , we have

$$\mathbf{i}^* \cdot \mathbf{v} = v_1^* \mathbf{i}^* \cdot \mathbf{i}^* + v_2^* \mathbf{i}^* \cdot \mathbf{j}^* + v_3^* \mathbf{i}^* \cdot \mathbf{k}^*.$$

Since the first scalar product on the right is 1 and the others are zero,

$$\mathbf{i}^* \cdot \mathbf{v} = v_1^*.$$

From this and (2) it follows that

$$v_1^* = \mathbf{i}^* \cdot i\mathbf{v}_1 + \mathbf{i}^* \cdot j\mathbf{v}_2 + \mathbf{i}^* \cdot k\mathbf{v}_3.$$

Similarly,

$$v_2^* = \mathbf{j}^* \cdot i\mathbf{v}_1 + \mathbf{j}^* \cdot j\mathbf{v}_2 + \mathbf{j}^* \cdot k\mathbf{v}_3,$$

$$v_3^* = \mathbf{k}^* \cdot i\mathbf{v}_1 + \mathbf{k}^* \cdot j\mathbf{v}_2 + \mathbf{k}^* \cdot k\mathbf{v}_3.$$

Hence the components of a vector \mathbf{v} in a Cartesian coordinate system can be expressed as linear functions of the components of \mathbf{v} with respect to another Cartesian coordinate system.

To write our transformation formulas in a simpler form we adopt the notation

$$(3) \quad \begin{aligned} \mathbf{i}^* \cdot \mathbf{i} &= a_{11} & \mathbf{i}^* \cdot \mathbf{j} &= a_{12} & \mathbf{i}^* \cdot \mathbf{k} &= a_{13} \\ \mathbf{j}^* \cdot \mathbf{i} &= a_{21} & \mathbf{j}^* \cdot \mathbf{j} &= a_{22} & \mathbf{j}^* \cdot \mathbf{k} &= a_{23} \\ \mathbf{k}^* \cdot \mathbf{i} &= a_{31} & \mathbf{k}^* \cdot \mathbf{j} &= a_{32} & \mathbf{k}^* \cdot \mathbf{k} &= a_{33}. \end{aligned}$$

Then we have

$$(4) \quad \begin{aligned} v_1^* &= a_{11}v_1 + a_{12}v_2 + a_{13}v_3 \\ v_2^* &= a_{21}v_1 + a_{22}v_2 + a_{23}v_3 \\ v_3^* &= a_{31}v_1 + a_{32}v_2 + a_{33}v_3. \end{aligned}$$

Using summation signs, we may write this more briefly

$$(4') \quad v_k^* = \sum_{l=1}^3 a_{kl} v_l, \quad k = 1, 2, 3.$$

A similar consideration leads to the inverse formulas

$$(5) \quad \begin{aligned} v_1 &= a_{11}v_1^* + a_{21}v_2^* + a_{31}v_3^* \\ v_2 &= a_{12}v_1^* + a_{22}v_2^* + a_{32}v_3^* \\ v_3 &= a_{13}v_1^* + a_{23}v_2^* + a_{33}v_3^* \end{aligned}$$

or, more briefly,

$$(5') \quad v_l = \sum_{m=1}^3 a_{ml} v_m^*, \quad l = 1, 2, 3.$$

Note that (4) and (5) contain the same coefficients a_{kl} , but these coefficients (except for a_{11}, a_{22}, a_{33}) occupy different positions in (4) and (5).

The geometrical interpretation of those nine coefficients a_{kl} is very simple. Since \mathbf{i} and \mathbf{i}^* are unit vectors, it follows from (1) in Sec. 5.4 that $a_{11} = \mathbf{i}^* \cdot \mathbf{i}$ is the cosine of the angle between the positive x^* and x axes. Similarly,

$a_{12} = \mathbf{i}^* \cdot \mathbf{j}$ is the cosine of the angle between the positive x^* and y axes, and so on.

The coefficients a_{kl} satisfy certain important relations which we shall now derive. By inserting (5') into (4') we find

$$(6) \quad v_k^* = \sum_{l=1}^3 a_{kl} v_l = \sum_{l=1}^3 a_{kl} \sum_{m=1}^3 a_{ml} v_m^* = \sum_{m=1}^3 v_m^* \left(\sum_{l=1}^3 a_{kl} a_{ml} \right),$$

where $k = 1, 2, 3$. For example, if $k = 1$, this becomes

$$v_1^* = v_1^* \left(\sum_{l=1}^3 a_{1l} a_{1l} \right) + v_2^* \left(\sum_{l=1}^3 a_{1l} a_{2l} \right) + v_3^* \left(\sum_{l=1}^3 a_{1l} a_{3l} \right).$$

In order that this relation hold for any vector $\mathbf{v} = v_1^* \mathbf{i}^* + v_2^* \mathbf{j}^* + v_3^* \mathbf{k}^*$, the first sum must be 1 and the other two sums must be zero. For $k = 2$ and $k = 3$ the situation is similar. Consequently, (6) holds for any vector if, and only if,

$$(7) \quad \sum_{l=1}^3 a_{kl} a_{ml} = \begin{cases} 0 & (k \neq m) \\ 1 & (k = m) \end{cases}.$$

Using the so-called *Kronecker*¹¹ symbol or Kronecker delta

$$\delta_{km} = \begin{cases} 0 & (k \neq m) \\ 1 & (k = m) \end{cases}$$

we may write (7) in the form

$$(7') \quad \sum_{l=1}^3 a_{kl} a_{ml} = \delta_{km} \quad (k, m = 1, 2, 3).$$

Forming three vectors with components

$$a_{11}, a_{12}, a_{13} \quad a_{21}, a_{22}, a_{23} \quad a_{31}, a_{32}, a_{33}$$

we see that the left side of (7') is the scalar product of two of these vectors, and (7') implies that these vectors are orthogonal unit vectors. Hence their scalar triple product has the value +1 or -1; that is,

$$(8) \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \pm 1.$$

We mention, without proof, that if both coordinate systems under consideration are right-handed (or both left-handed), then the determinant has the value +1, whereas if one system is right-handed and the other left-handed, the determinant has the value -1. We may now sum up our result as follows.

¹¹ LEOPOLD KRONECKER (1823–1891), German mathematician, who made important contributions to algebra and the theory of numbers.

Theorem 1. The components v_1, v_2, v_3 and v_1^*, v_2^*, v_3^* of any vector \mathbf{v} with respect to any two Cartesian coordinate systems can be obtained from each other by means of (4) and (5), where the coefficients a_{kl} are given by (3) and satisfy (7) and (8).

From our consideration of vector components we may now immediately obtain the formulas for the transformation of a Cartesian coordinate system into any other such system as follows.

If the xyz and $x^*y^*z^*$ coordinate systems have the same origin, then \mathbf{v} may be bound at the origin and regarded as the position vector of its terminal point Q . If (x, y, z) and (x^*, y^*, z^*) are the coordinates of Q with respect to the two coordinate systems, then in (4) and (5)

$$v_1 = x, v_2 = y, v_3 = z \quad \text{and} \quad v_1^* = x^*, v_2^* = y^*, v_3^* = z^*.$$

Consequently, (4) and (5), written in terms of x, y, z and x^*, y^*, z^* instead of v_1, v_2, v_3 and v_1^*, v_2^*, v_3^* , represent the transformations between those two coordinate systems with common origin.

The most general transformation of a Cartesian coordinate system into another such system may be decomposed into a transformation of the type just considered and a translation. Under a translation, corresponding coordinates differ merely by a constant. We thus obtain

Theorem 2. The transformation of any Cartesian xyz -coordinate system into any other Cartesian $x^*y^*z^*$ -coordinate system is of the form

$$(9) \quad \begin{aligned} x^* &= a_{11}x + a_{12}y + a_{13}z + b_1 \\ y^* &= a_{21}x + a_{22}y + a_{23}z + b_2 \\ z^* &= a_{31}x + a_{32}y + a_{33}z + b_3 \end{aligned}$$

and conversely

$$(10) \quad \begin{aligned} x &= a_{11}x^* + a_{21}y^* + a_{31}z^* + c_1 \\ y &= a_{12}x^* + a_{22}y^* + a_{32}z^* + c_2 \\ z &= a_{13}x^* + a_{23}y^* + a_{33}z^* + c_3 \end{aligned}$$

where the coefficients a_{kl} are given by (3) and satisfy (7) and (8) and $b_1, b_2, b_3, c_1, c_2, c_3$ are constants.

Important applications of our results will be considered in the following two sections.

Problems

- Interpret all the coefficients in (4') geometrically.

Determine the constants a_{kl} and b_k such that (9) represents:

- A translation which carries the origin into the point $(1, 3, 2)$
- A translation which carries the point $(2, -1, 0)$ into the point $(1, 4, 2)$

4. A reflection in the xy -plane
5. A reflection in the plane $y = z$
6. A rotation about the x -axis through an angle θ
7. A rotation such that the positive x^*, y^*, z^* axes coincide with the positive y, z, x axes, respectively
8. What are the values of the determinant (8) in Probs. 2-7?
9. Derive (5).
10. From (1), Sec. 5.4, it follows that the value of a scalar product is independent of the particular choice of Cartesian coordinates. Using (4) or (5) in this section, show that this follows also from (8), Sec. 5.4.

5.18 Divergence of a Vector Field

Let $\mathbf{v}(x, y, z)$ be a differentiable vector function, where x, y, z are Cartesian coordinates in space, and let v_1, v_2, v_3 be the components of \mathbf{v} . Then the function

$$(1) \quad \operatorname{div} \mathbf{v} = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}$$

is called the **divergence of \mathbf{v}** or the *divergence of the vector field defined by \mathbf{v}* .

Another common notation for the divergence of \mathbf{v} is $\nabla \cdot \mathbf{v}$,

$$\begin{aligned} \operatorname{div} \mathbf{v} &= \nabla \cdot \mathbf{v} = \left(\frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial y} \mathbf{j} + \frac{\partial}{\partial z} \mathbf{k} \right) \cdot (v_1 \mathbf{i} + v_2 \mathbf{j} + v_3 \mathbf{k}) \\ &= \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}, \end{aligned}$$

with the understanding that the “product” $(\partial/\partial x)v_1$ in the scalar product means the partial derivative $\partial v_1/\partial x$, etc. This is a convenient notation, but nothing more. Note that $\nabla \cdot \mathbf{v}$ means the scalar $\operatorname{div} \mathbf{v}$ whereas ∇f means the vector $\operatorname{grad} f$ defined in Sec. 5.16.

We shall see later that the divergence has an important physical meaning. Clearly the values of a function which characterizes a physical or geometrical property must be independent of the particular choice of coordinates; that is, those values must be invariant with respect to coordinate transformations.

Theorem 1. *The divergence is a scalar function; that is, the values of $\operatorname{div} \mathbf{v}$ depend only on the points in space (and, of course, on \mathbf{v}) but not on the particular choice of the coordinates in (1), so that with respect to other Cartesian coordinates x^*, y^*, z^* and corresponding components v_1^*, v_2^*, v_3^* of \mathbf{v} the function $\operatorname{div} \mathbf{v}$ is given by*

$$(2) \quad \operatorname{div} \mathbf{v} = \frac{\partial v_1^*}{\partial x^*} + \frac{\partial v_2^*}{\partial y^*} + \frac{\partial v_3^*}{\partial z^*}.$$

Proof. We shall derive (2) from (1). To simplify our formulas in this proof, we adopt the notation

$$x_1 = x, \quad x_2 = y, \quad x_3 = z \quad \text{and} \quad x_1^* = x^*, \quad x_2^* = y^*, \quad x_3^* = z^*.$$

Using this notation, we may write formula (9) in Sec. 5.17 in the short form

$$(3) \quad x_k^* = \sum_{l=1}^3 a_{kl} x_l + b_k \quad (k = 1, 2, 3).$$

By applying the chain rule for functions of several variables (cf. Sec. 5.15) we obtain

$$(4) \quad \frac{\partial v_l}{\partial x_i} = \sum_{k=1}^3 \frac{\partial v_l}{\partial x_k^*} \frac{\partial x_k^*}{\partial x_i}.$$

In this sum, $\partial x_k^*/\partial x_i = a_{ki}$, as follows from (3). Furthermore, by (5') in Sec. 5.17,

$$v_l = \sum_{m=1}^3 a_{ml} v_m^*.$$

By differentiating this we have

$$\frac{\partial v_l}{\partial x_k^*} = \sum_{m=1}^3 a_{ml} \frac{\partial v_m^*}{\partial x_k^*}.$$

If we substitute these expressions into (4), we obtain

$$\frac{\partial v_l}{\partial x_i} = \sum_{k=1}^3 \sum_{m=1}^3 a_{ml} \frac{\partial v_m^*}{\partial x_k^*} a_{ki} \quad l = 1, 2, 3.$$

By adding these three formulas we find

$$\operatorname{div} \mathbf{v} = \sum_{i=1}^3 \frac{\partial v_i}{\partial x_i} = \sum_{k=1}^3 \sum_{m=1}^3 \sum_{l=1}^3 a_{kl} a_{ml} \frac{\partial v_m^*}{\partial x_k^*}.$$

Because of (7') in Sec. 5.17 this reduces to

$$\operatorname{div} \mathbf{v} = \sum_{k=1}^3 \sum_{m=1}^3 \delta_{km} \frac{\partial v_m^*}{\partial x_k^*} = \frac{\partial v_1^*}{\partial x_1^*} + \frac{\partial v_2^*}{\partial x_2^*} + \frac{\partial v_3^*}{\partial x_3^*}.$$

We see that the expression on the right is identical with that in (2), and the proof is complete.

If $f(x, y, z)$ is a twice differentiable scalar function, then

$$\operatorname{grad} f = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} + \frac{\partial f}{\partial z} \mathbf{k}$$

and by (1),

$$\operatorname{div} (\operatorname{grad} f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$

The expression on the right is the Laplacian of f (cf. Sec. 5.16). Thus

$$(5) \quad \operatorname{div}(\operatorname{grad} f) = \nabla^2 f.$$

Example 1. The gravitational force \mathbf{p} in Ex. 4, Sec. 5.16, is the gradient of the scalar function $f(x, y, z) = c/r$, which satisfies Laplace's equation $\nabla^2 f = 0$. According to (5), this means that $\operatorname{div} \mathbf{p} = 0$ ($r > 0$).

The following example, taken from hydrodynamics, may serve as an introductory illustration of the physical significance of the divergence of a vector field. A more detailed physical interpretation of the divergence will be given later (in Sec. 6.9).

Example 2 (Motion of a compressible fluid). We consider the motion of a fluid in a region R having no sources or sinks in R , that is, no points at which fluid is produced or disappears. The concept of *fluid state* is meant to cover also gases and vapors. Fluids in the restricted sense, or *liquids*, have small compressibility which can be neglected in many problems, whereas gases and vapors have considerable compressibility; that is, their density ρ (= mass per unit volume) depends on the coordinates x, y, z in space (and may depend on the time t). We assume that our fluid is compressible.

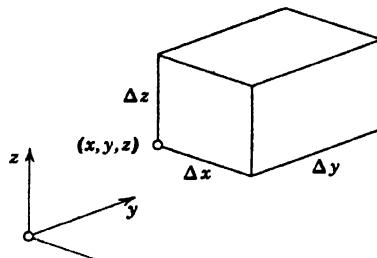


Fig. 169. Physical interpretation of the divergence.

We consider the flow through a small rectangular parallelepiped W of dimensions¹² $\Delta x, \Delta y, \Delta z$ with edges parallel to the coordinate axes (Fig. 169). W has the volume $\Delta V = \Delta x \Delta y \Delta z$. Let

$$\mathbf{v} = v_1 \mathbf{i} + v_2 \mathbf{j} + v_3 \mathbf{k}$$

be the velocity vector of the motion. We set

$$(6) \quad \mathbf{u} = \rho \mathbf{v} = u_1 \mathbf{i} + u_2 \mathbf{j} + u_3 \mathbf{k}$$

and assume that \mathbf{u} and \mathbf{v} are continuously differentiable vector functions of x, y, z , and t . Let us calculate the change in the mass included in W by considering the flux across the boundary, that is, the total loss of mass leaving W per unit time. Consider the flow through the left-hand face of W , whose area is $\Delta x \Delta z$. The components v_1 and v_3 of \mathbf{v} are parallel to that face and contribute nothing to that flow. Hence the mass of fluid entering through that face during a short time interval Δt is given approximately by

$$(\rho v_2)_y \Delta x \Delta z \Delta t = u_2(y) \Delta x \Delta z \Delta t$$

¹² It is a standard usage to indicate small quantities by Δ ; this has, of course, nothing to do with the Laplacian.

where the subscript y indicates that this expression refers to the left-hand face. The mass of fluid leaving W through the opposite face during the same time interval is approximately $u_2(y + \Delta y) \Delta x \Delta z \Delta t$. The difference

$$\Delta u_2 \Delta x \Delta z \Delta t = \frac{\Delta u_2}{\Delta y} \Delta V \Delta t \quad [\Delta u_2 = u_2(y + \Delta y) - u_2(y)]$$

is the approximate loss of mass. Two similar expressions are obtained by considering the other two pairs of parallel faces of W . If we add these three expressions, we find that the total loss of mass in W during the time interval Δt is approximately

$$\left(\frac{\Delta u_1}{\Delta x} + \frac{\Delta u_2}{\Delta y} + \frac{\Delta u_3}{\Delta z} \right) \Delta V \Delta t$$

where

$$\Delta u_1 = u_1(x + \Delta x) - u_1(x) \quad \text{and} \quad \Delta u_3 = u_3(z + \Delta z) - u_3(z).$$

This loss of mass in W is caused by the time rate of change of the density and is thus equal to

$$-\frac{\partial \rho}{\partial t} \Delta V \Delta t.$$

If we equate both expressions, divide the resulting equation by $\Delta V \Delta t$, and let Δx , Δy , Δz , and Δt approach zero, then we obtain

$$\operatorname{div} \mathbf{u} = \operatorname{div} (\rho \mathbf{v}) = -\frac{\partial \rho}{\partial t}$$

or

$$(7) \quad \frac{\partial \rho}{\partial t} + \operatorname{div} (\rho \mathbf{v}) = 0.$$

This important relation is called the *condition for the conservation of mass* or the *continuity equation of a compressible fluid flow*.

If the flow is *steady*, that is, independent of time t , then $\partial \rho / \partial t = 0$, and the continuity equation is

$$(8) \quad \operatorname{div} (\rho \mathbf{v}) = 0.$$

If the fluid is *incompressible*, then the density ρ is constant, and (8) becomes

$$(9) \quad \operatorname{div} \mathbf{v} = 0. \checkmark$$

This relation is known as the *condition of incompressibility*. It expresses the fact that the balance of outflow and inflow for a given volume element is zero at any time. Clearly the assumption that the flow has no sources or sinks in R is essential to our argument.

Problems

Find the divergence of the following vector functions.

- | | |
|--|--|
| 1. $x\mathbf{i} + y\mathbf{j}$ | 2. $y^2\mathbf{i} + x^2\mathbf{j}$ |
| 3. $(x^2 - 3y^2)\mathbf{i}$ | 4. $x^2\mathbf{i} + y^2\mathbf{j} - zk$ |
| 5. $xyz(\mathbf{i} + \mathbf{j} + \mathbf{k})$ | 6. $x^2\mathbf{i} + y^2\mathbf{j} + z^2\mathbf{k}$ |
| 7. $yz^2\mathbf{i} - zx^2\mathbf{k}$ | 8. $(xi + yj + zk)/(x^2 + y^2 + z^2)^{3/2}$ |

9. It is obvious from (1) that $\operatorname{div} \mathbf{v}$ is a scalar?

Show that

- | | |
|---|--|
| 10. $\operatorname{div} (\operatorname{grad} f) = \nabla^2 f$ | 11. $\operatorname{div} (f\mathbf{v}) = f \operatorname{div} \mathbf{v} + \mathbf{v} \cdot \nabla f$ |
| 12. $\operatorname{div} (f\nabla g) = f \nabla^2 g + \nabla f \cdot \nabla g$ | 13. $\operatorname{div} (f\nabla g) - \operatorname{div}(g\nabla f) = f \nabla^2 g - g \nabla^2 f$ |

14. Show that the continuity equation (7) may be written

$$\frac{\partial \rho}{\partial t} + \rho \operatorname{div} \mathbf{v} + \mathbf{v} \cdot \nabla \rho = 0.$$

Using the formula in Prob. 12, find the divergence of

15. $e^x(\sin y \mathbf{i} + \cos y \mathbf{j})$

16. $(xi + yj + zk)/(x^2 + y^2 + z^2)^{3/2}$

17. Consider a steady flow whose velocity vector is $\mathbf{v} = yi$. Show that it has the following properties. The flow is incompressible. The particles which at time $t = 0$ are in the cube bounded by the planes $x = 0, x = 1, y = 0, y = 1, z = 0, z = 1$ occupy at $t = 1$ the volume 1.

18. Consider a steady flow having the velocity $\mathbf{v} = xi$. Show that the individual particles have the position vectors $\mathbf{r}(t) = c_1 e^t \mathbf{i} + c_2 \mathbf{j} + c_3 \mathbf{k}$, where c_1, c_2, c_3 are constants, the flow is compressible, and the particles which at $t = 0$ fill the cube in Prob. 17 occupy at $t = 1$ the volume e .

Find the directional derivative of $\operatorname{div} \mathbf{u}$ at the point $P(4, 4, 2)$ in the direction of the corresponding outer normal of the sphere $x^2 + y^2 + z^2 = 36$ where

19. $\mathbf{u} = x^4 \mathbf{i} + y^4 \mathbf{j} + z^4 \mathbf{k}$

20. $\mathbf{u} = xz \mathbf{i} + yx \mathbf{j} + zy \mathbf{k}$

5.19 Curl of a Vector Field

Let x, y, z be right-handed Cartesian coordinates in space, and let

$$\mathbf{v}(x, y, z) = v_1 \mathbf{i} + v_2 \mathbf{j} + v_3 \mathbf{k}$$

be a differentiable vector function. Then the function

$$(1) \quad \begin{aligned} \operatorname{curl} \mathbf{v} = \nabla \times \mathbf{v} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ v_1 & v_2 & v_3 \end{vmatrix} \\ &= \left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z} \right) \mathbf{i} + \left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x} \right) \mathbf{j} + \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right) \mathbf{k} \end{aligned}$$

is called the *curl of the vector function v* or the *curl of the vector field defined by v*. In the case of a left-handed Cartesian coordinate system, the symbolic determinant in (1) is preceded by a minus sign, in agreement with (5) in Sec. 5.6.

Instead of $\operatorname{curl} \mathbf{v}$ the notation $\operatorname{rot} \mathbf{v}$ is also used in the literature.

Theorem 1. *Curl v is a vector.*

The proof of this theorem is included in the answer to Prob. 19.

The curl plays an important role in many applications. Its significance will be explained in more detail in Sec. 6.11. Now we confine ourselves to some simple examples and remarks.

Example 1 (Rotation of a rigid body). We have seen in Ex. 3, Sec. 5.6, that a rotation of a rigid body B about a fixed axis in space can be described by a vector w of magnitude ω in the direction of the axis of rotation, where $\omega (>0)$ is the angular speed of the rotation, and w is directed so that the rotation appears clockwise if we look in the direction of w . According to (7), Sec. 5.6, the velocity field of the rotation can be represented in the form

$$\mathbf{v} = \mathbf{w} \times \mathbf{r}$$

where \mathbf{r} is the position vector of a moving point with respect to a Cartesian coordinate system having the origin on the axis of rotation. Let us choose right-handed Cartesian coordinates such that $w = \omega k$; that is, the axis of rotation is the z -axis. Then (cf. Ex. 3 in Sec. 5.9)

$$\mathbf{v} = \mathbf{w} \times \mathbf{r} = -\omega y \mathbf{i} + \omega x \mathbf{j}$$

and, therefore,

$$\operatorname{curl} \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -\omega y & \omega x & 0 \end{vmatrix} = 2\omega \mathbf{k},$$

that is,

$$(2) \quad \operatorname{curl} \mathbf{v} = 2\omega \mathbf{k}.$$

Hence, in the case of a rotation of a rigid body, the curl of the velocity field has the direction of the axis of rotation, and its magnitude equals twice the angular speed ω of the rotation.

Note that our result does not depend on the particular choice of the Cartesian coordinate system, since we are dealing with vectors.

For any twice continuously differentiable scalar function f ,

$$(3) \quad \operatorname{curl} (\operatorname{grad} f) = 0, \quad \checkmark$$

as can easily be verified by direct calculation. In words: if a vector function is the gradient of a scalar function, its curl is the zero vector. Since the curl characterizes the rotation in a field, we also say more briefly that gradient fields describing a motion are irrotational. (If such a field occurs in some other connection, not as a velocity field, it is usually called conservative; cf. at the end of Sec. 5.16.)

Example 2. The gravitational field in Ex. 4, Sec. 5.16, has $\operatorname{curl} \mathbf{p} = 0$. The field in Ex. 1, this section, is not irrotational. A similar velocity field is obtained by stirring coffee in a cup.

Problems

Find $\operatorname{curl} \mathbf{v}$ where, with respect to right-handed Cartesian coordinates,

1. $\mathbf{v} = y \mathbf{i} - x \mathbf{j}$
2. $\mathbf{v} = y \mathbf{i} + z \mathbf{j} + x \mathbf{k}$
3. $\mathbf{v} = z^2 \mathbf{i} + x^2 \mathbf{j} + y^2 \mathbf{k}$
4. $\mathbf{v} = x^2 \mathbf{i} + y^2 \mathbf{j} + z^2 \mathbf{k}$
5. $\mathbf{v} = xz \mathbf{i} - yz \mathbf{j}$
6. $\mathbf{v} = (xi + yj + zk)/(x^2 + y^2 + z^2)^{3/2}$
7. $\mathbf{v} = e^{xyz}(i + j + k)$
8. $\mathbf{v} = e^x \sin y \mathbf{i} + e^x \cos y \mathbf{j}$
9. A steady fluid motion has velocity $\mathbf{v} = xi + yj$. Find $\operatorname{curl} \mathbf{v}$. Is the motion incompressible? Find the paths of the particles.

10. Same situation and questions as in Prob. 9 when $\mathbf{v} = y\mathbf{i} - z\mathbf{j}$.
11. Same situation and questions as in Prob. 9 when $\mathbf{v} = y^2\mathbf{i}$.
12. Verify by direct calculation that, in Ex. 2, $\operatorname{curl} \mathbf{p} = 0$.

Show that, granted sufficient differentiability,

- | | |
|--|--|
| 13. $\operatorname{curl}(\mathbf{u} + \mathbf{v}) = \operatorname{curl} \mathbf{u} + \operatorname{curl} \mathbf{v}$ | 14. $\operatorname{curl}(f\mathbf{v}) = \operatorname{grad} f \times \mathbf{v} + f \operatorname{curl} \mathbf{v}$ |
| 15. $\operatorname{div}(\operatorname{curl} \mathbf{v}) = 0$ | 16. $\operatorname{div}(\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot \operatorname{curl} \mathbf{u} - \mathbf{u} \cdot \operatorname{curl} \mathbf{v}$ |

With respect to right-handed Cartesian coordinates, let $\mathbf{u} = y\mathbf{i} + z\mathbf{j} + z\mathbf{k}$ and $\mathbf{v} = xy\mathbf{i} + yz\mathbf{j} + zz\mathbf{k}$. Find

17. $\operatorname{curl}(\mathbf{u} \times \mathbf{v}), \operatorname{div}(\mathbf{u} \times \mathbf{v})$
18. $\mathbf{u} \times \operatorname{curl} \mathbf{v}, \mathbf{v} \times \operatorname{curl} \mathbf{u}$
19. Prove Theorem 1.
20. Using Probs. 14 and 16, show that $\operatorname{div}(g \nabla f \times f \nabla g) = 0$.

Chapter 6

Line and Surface Integrals. Integral Theorems

In this chapter we shall define line integrals and surface integrals, and consider some important applications of such integrals which occur frequently in connection with physical and engineering problems. We shall see that a line integral is a natural generalization of a definite integral, and a surface integral is a generalization of a double integral.

Line integrals can be transformed into double integrals (Sec. 6.4) or into surface integrals (Sec. 6.10), and conversely. Triple integrals can be transformed into surface integrals (Sec. 6.8). These transformations are of great practical importance. The corresponding formulas of Gauss, Green, and Stokes serve as powerful tools in many applications as well as in theoretical problems. We shall see that they also lead to a better understanding of the physical meaning of the divergence and the curl of a vector function.

Prerequisites for this chapter: elementary integral calculus, vector analysis (Chap. 5).

Sections which may be omitted in a shorter course: 6.6, 6.9, 6.11.

References: Appendix 1, Part C.

Answers to problems: Appendix 2.

6.1 Line Integral

The concept of a line integral is a simple and natural generalization of the concept of a definite integral

$$(1) \quad \int_a^b f(x) dx.$$

In (1) we integrate along the x -axis, and the integrand $f(x)$ is a (single-valued) function defined at each point between a and b . We shall see that, in the case of a line integral, we integrate along a curve in space (or in the plane), and the integrand is a single-valued function defined at the points of that curve.

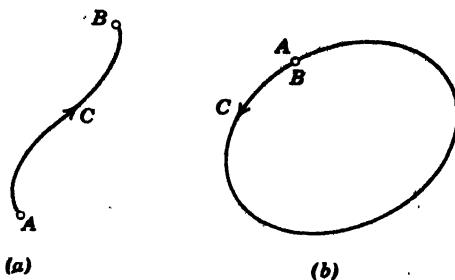


Fig. 170. Oriented curve.

The way of defining a line integral is quite similar to the familiar way of defining a definite integral known from calculus. We may proceed as follows.

We consider a curve C in space and orient C by choosing one of the two directions along C as the *positive direction*; the opposite direction along C is then called the *negative direction*. Let A be the initial point and B the terminal point of C under the chosen orientation. (These points may coincide as in Fig. 170b; then C is closed.)

We may now represent C by a parametric representation

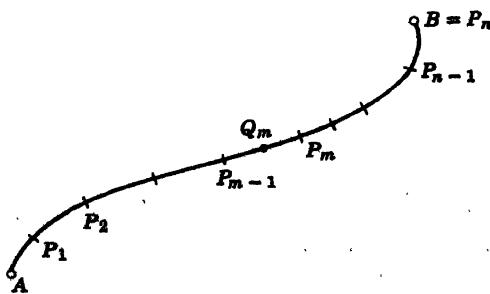
$$(2) \quad \mathbf{r}(s) = x(s)\mathbf{i} + y(s)\mathbf{j} + z(s)\mathbf{k} \quad (a \leq s \leq b),$$

where s is the arc length of C (cf. Sec. 5.12) and A and B correspond to $s = a$ and $s = b$, respectively.

We assume that $\mathbf{r}(s)$ is continuous and has a continuous first derivative which is different from the zero vector for all s under consideration. Then C has a unique tangent at each of its points. A curve satisfying these assumptions will be called a *smooth curve*.

Let $f(x, y, z)$ be a given function which is defined (at least) at each point of C , and is a continuous function of s . Instead of $f[x(s), y(s), z(s)]$ we simply write $f(P)$ where P is the point on C with coordinates $x(s), y(s), z(s)$. We subdivide C into n portions in an arbitrary manner (Fig. 171); let $P_0 (= A)$, $P_1, \dots, P_{n-1}, P_n (= B)$ be the end points of these portions, and let

$$s_0 (= a) < s_1 < s_2 < \dots < s_n (= b)$$

Fig. 171. Subdivision of C .

be the corresponding values of s . Then we choose an arbitrary point on each portion, say, a point Q_1 between P_0 and P_1 , a point Q_2 between P_1 and P_2 , etc., and form the sum

$$J_n = \sum_{m=1}^n f(Q_m) \Delta s_m$$

where

$$\Delta s_m = s_m - s_{m-1}.$$

This we do for each $n = 2, 3, \dots$ in a completely independent manner, but such that the greatest Δs_m approaches zero as n approaches infinity. In this way we obtain a sequence of real numbers J_2, J_3, \dots . The limit of this sequence is called the **line integral of f along C from A to B** , and is denoted by

$$\int_C f(P) ds \quad \text{or} \quad \int_C f(x, y, z) ds.$$

The curve C is called the *path of integration*.

Since, by assumption, f is continuous and C is smooth, that limit exists and is independent of the choice of those subdivisions and points Q_m . In fact, the position of a point P on C is determined by the corresponding value of the arc length s ; since A and B correspond to $s = a$ and $s = b$, respectively, we thus have

$$\int_C f(P) ds = \int_a^b f(s) ds.$$

Hence the line integral is equal to the definite integral on the right, for which the statement holds, as is known from calculus.

Throughout this book all paths of integration for line integrals will be assumed to be piecewise smooth, that is, will consist of finitely many smooth curves.

From the definition it follows that familiar properties of ordinary definite integrals are equally valid for line integrals:

- $$(3)$$
- (a) $\int_C kf ds = k \int_C f ds$ (k constant)
 - (b) $\int_C (f + g) ds = \int_C f ds + \int_C g ds$
 - (c) $\int_C f ds = \int_{C_1} f ds + \int_{C_2} f ds$

where in (3c) the path C is subdivided into two arcs C_1 and C_2 which have the same orientation as C (Fig. 172). In (3b) the orientation of C is the same in all three integrals.

If the sense of integration along C is reversed, the value of the integral is multiplied by -1 .

Examples will be considered in the following sections.

Fig. 172. Formula (3c).

6.2 Evaluation of Line Integrals

To evaluate a line integral we first have to write it as an ordinary definite integral. This is quite simple and can be done by using the given representation of the path of integration C .

If C is represented in the form

$$\mathbf{r}(s) = x(s)\mathbf{i} + y(s)\mathbf{j} + z(s)\mathbf{k} \quad (a \leq s \leq b),$$

[cf. (2) in Sec. 6.1], then we immediately have

$$\int_C f(x, y, z) ds = \int_a^b f[x(s), y(s), z(s)] ds,$$

the integral on the right being a definite integral.

In most cases the representation of C will be of the form

$$(1) \quad \mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k} \quad (t_0 \leq t \leq t_1)$$

or may easily be converted to this form; here, t denotes any parameter, and $\mathbf{r}(t)$ is continuous and has a continuous derivative. Then, using the familiar substitution formula for definite integrals, we obtain

$$\int_a^b f[x(s), y(s), z(s)] ds = \int_{t_0}^{t_1} f[x(t), y(t), z(t)] \frac{ds}{dt} dt.$$

Therefore, if C is represented in the form (1), then

$$(2) \quad \int_C f(x, y, z) ds = \int_{t_0}^{t_1} f[x(t), y(t), z(t)] \frac{ds}{dt} dt$$

where, according to (3) in Sec. 5.12,

$$\frac{ds}{dt} = \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}.$$

Example 1. Evaluate $\int_C xy^3 ds$, where C is the segment of the line $y = 2x$ in the xy -plane from $A: (-1, -2, 0)$ to $B: (1, 2, 0)$.

Representing C in the form¹

$$\mathbf{r}(t) = t\mathbf{i} + 2t\mathbf{j} \quad (-1 \leq t \leq 1)$$

we have

$$\dot{\mathbf{r}} = \mathbf{i} + 2\mathbf{j} \quad \text{and} \quad \frac{ds}{dt} = \sqrt{\dot{x}^2 + \dot{y}^2} = \sqrt{5}.$$

On C , $xy^3 = t(2t)^3 = 8t^4$, and therefore

$$\int_C xy^3 ds = 8\sqrt{5} \int_{-1}^1 t^4 dt = \frac{16}{5} \approx 7.16.$$

¹ Of course, since $x = t$ we may write x in place of t .

Example 2. Evaluate $\int_C (x^2 + y^2 + z^2)^2 ds$ where C is the arc of the circular helix (Sec. 5.11)

$$\mathbf{r}(t) = \cos t \mathbf{i} + \sin t \mathbf{j} + 3t \mathbf{k}$$

from $A: (1, 0, 0)$ to $B: (1, 0, 6\pi)$. We find $ds/dt = \sqrt{10}$. On C ,

$$(x^2 + y^2 + z^2)^2 = [\cos^2 t + \sin^2 t + (3t)^2]^2 = (1 + 9t^2)^2.$$

Since C corresponds to $0 \leq t \leq 2\pi$, we thus have

$$\begin{aligned} \int_C (x^2 + y^2 + z^2)^2 ds &= \sqrt{10} \int_0^{2\pi} (1 + 9t^2)^2 dt \\ &= \sqrt{10} [2\pi + 6(2\pi)^3 + \frac{81}{5}(2\pi)^5] \approx 506\,400. \end{aligned}$$

Line integrals which involve empirically given functions or lead to complicated definite integrals may be evaluated by using numerical or graphical methods of integration.

In many applications the integrands of line integrals are of the form

$$(3) \quad g(x, y, z) \frac{dx}{ds}, \quad g(x, y, z) \frac{dy}{ds}, \quad \text{or} \quad g(x, y, z) \frac{dz}{ds},$$

where dx/ds , dy/ds , and dz/ds are the derivatives of the functions occurring in the parametric representation of the path of integration. Then we simply write

$$(4) \quad \int_C g(x, y, z) \frac{dx}{ds} ds = \int_C g(x, y, z) dx,$$

and similarly in the other two cases. For sums of these types of integrals along the same path C we adopt the simplified notation

$$(5) \quad \int_C f dx + \int_C g dy + \int_C h dz = \int_C (f dx + g dy + h dz).$$

Using the representation of C , we may eliminate two of the three independent variables in the integrand and then evaluate the resulting definite integral in which the remaining independent variable is the variable of integration.

Example 3. Evaluate the line integral $I = \int_C [x^3 y dx + (x - z) dy + xyz dz]$ where C is the arc of the parabola $y = x^2$ in the plane $z = 2$ from $A: (0, 0, 2)$ to $B: (1, 1, 2)$. Since $y = x^2$, we have $dy/dx = 2x$ or $dy = 2x dx$. Since $z = 2$ is constant, it follows that the integral of the last term in the integrand is zero. Thus

$$\int_0^1 [x^3 x^2 dx + (x - 2) 2x dx] = \int_0^1 (x^6 + 2x^3 - 4x) dx = -\frac{17}{15}.$$

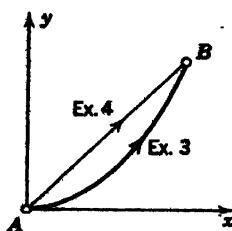


Fig. 173. Paths in Examples 3 and 4.

Example 4. Evaluate the line integral I in Ex. 3 where C is the segment of the straight line $y = x$, $z = 2$ from $A: (0, 0, 2)$ to $B: (1, 1, 2)$. Now $dy = dx$, and

$$I = \int_0^1 (x^3 + x - 2) dx = -\frac{5}{4}.$$

In Ex. 3 and Ex. 4 the integrands and the end points of the paths are the same, but the values of I are different. This illustrates the important fact that, in general, the value of a line integral of a given function depends not only on the end points but also on the geometric shape of the path of integration. We shall consider this basic fact in Sec. 6.12.

In many cases, the functions f , g , and h in (5) are the components v_1 , v_2 , and v_3 , respectively, of a vector function

$$\mathbf{v} = v_1 \mathbf{i} + v_2 \mathbf{j} + v_3 \mathbf{k}.$$

Then

$$v_1 dx + v_2 dy + v_3 dz = \left(v_1 \frac{dx}{ds} + v_2 \frac{dy}{ds} + v_3 \frac{dz}{ds} \right) ds,$$

the expression in parentheses being the scalar product of the vector \mathbf{v} and the unit tangent vector

$$\frac{d\mathbf{r}}{ds} = \frac{dx}{ds} \mathbf{i} + \frac{dy}{ds} \mathbf{j} + \frac{dz}{ds} \mathbf{k} \quad (\text{cf. Sec. 5.13}),$$

where $\mathbf{r}(s)$ represents the path of integration C . Therefore,

$$(6) \quad \int_C (v_1 dx + v_2 dy + v_3 dz) = \int_C \mathbf{v} \cdot \frac{d\mathbf{r}}{ds} ds.$$

This is sometimes written

$$\int_C \mathbf{v} \cdot \frac{d\mathbf{r}}{ds} ds = \int_C \mathbf{v} \cdot d\mathbf{r}$$

where

$$d\mathbf{r} = dx \mathbf{i} + dy \mathbf{j} + dz \mathbf{k},$$

but we should note that this is merely a convenient notation.

Example 5 (Work done by a force). Consider a particle on which a variable force \mathbf{p} acts. Let the particle be displaced along a given path C in space. Then the work W done by \mathbf{p} in this displacement is given by the line integral

$$(7) \quad W = \int_C \mathbf{p} \cdot d\mathbf{r},$$

the integration being taken in the sense of the displacement. This definition of work is suggested by that in Ex. 1, Sec. 5.4, together with the definition of a line integral as the limit of a sum.

We may introduce the time t as the variable of integration. Then

$$d\mathbf{r} = \frac{d\mathbf{r}}{dt} dt = \mathbf{v} dt$$

where \mathbf{v} is the velocity vector (cf. Sec. 5.14). Hence the line integral (7) becomes

$$(7') \quad W = \int_{t_0}^{t_1} \mathbf{p} \cdot \mathbf{v} dt$$

where t_0 and t_1 are the initial and final values of t .

By Newton's second law (Sec. 2.6),

$$\mathbf{p} = m\ddot{\mathbf{r}} = m\dot{\mathbf{v}},$$

where m is the mass of the particle. By inserting this into (7') we obtain

$$W = \int_{t_0}^{t_1} m\dot{\mathbf{v}} \cdot \mathbf{v} dt = \int_{t_0}^{t_1} \frac{d}{dt} \left(\frac{m}{2} \mathbf{v} \cdot \mathbf{v} \right) dt = \int_{t_0}^{t_1} \frac{d}{dt} \left(\frac{m}{2} |\mathbf{v}|^2 \right) dt = \frac{m}{2} |\mathbf{v}|^2 \Big|_{t_0}^{t_1};$$

that is, *the work done equals the gain in kinetic energy*. This is a basic law in mechanics.

We finally mention that for a line integral over a *closed* path C the symbol

$\left(\text{instead of } \int_C \right)$

is sometimes used in the literature.

Problems

Orienting C so that the sense of integration becomes the positive sense on C , evaluate $\int_C (x^2 + y^2) ds$:

1. Over the path $y = 2x$ from $(0, 0)$ to $(1, 2)$
2. Over the path $y = -x$ from $(1, -1)$ to $(2, -2)$
3. Counterclockwise along the circle $x^2 + y^2 = 4$ from $(2, 0)$ to $(0, 2)$
4. Counterclockwise around the circle $x^2 + y^2 = 1$ from $(1, 0)$ to $(1, 0)$
5. Over the x -axis from $(0, 0)$ to $(1, 0)$ and then parallel to the y -axis from $(1, 0)$ to $(1, 1)$
6. Over the y -axis from $(0, 0)$ to $(0, 1)$ and then parallel to the x -axis from $(0, 1)$ to $(1, 1)$
7. Evaluate $\int_C x^{-1}(y + z) ds$ where C is the arc of the circle $x^2 + y^2 = 4$, $z = 0$, from $(2, 0, 0)$ to $(\sqrt{2}, \sqrt{2}, 0)$ (counterclockwise).

Evaluate $\int_C (y^3 dx - 2x^3 dy)$:

8. Along the straight line from $(0, 2)$ to $(1, 1)$
9. Along the parabola $y = x^2$ from $(0, 0)$ to $(2, 4)$
10. Counterclockwise along the circle $x^2 + y^2 = 1$ from $(1, 0)$ to $(0, 1)$

Find the work done by the force $\mathbf{p} = xi - zj + 2yk$ in the displacement:

11. Along the y -axis from 0 to 1
12. Along the parabola $y = 2x^2$, $z = 2$ from $(0, 0, 2)$ to $(1, 2, 2)$
13. Along the curve $z = y^4$, $x = 1$ from $(1, 0, 0)$ to $(1, 1, 1)$
14. Along the closed path in Fig. 174

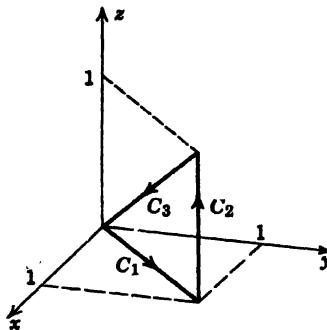


Fig. 174. Problem 14.

15. Along the straight line $y = x$, $z = x$ from $(1, 1, 1)$ to $(3, 3, 3)$
16. Along the curve $y = x$, $z = x^2$ from $(0, 0, 0)$ to $(1, 1, 1)$
17. Along the curve $y = x^3$, $z = 2$ from $(1, 1, 2)$ to $(2, 8, 2)$
18. Let \mathbf{p} be a vector function defined at all points of a curve C , and suppose that $|\mathbf{p}|$ is bounded, say, $|\mathbf{p}| < M$ on C , where M is some positive number. Show that

$$(8) \quad \left| \int_C \mathbf{p} \cdot d\mathbf{r} \right| < Mi$$

where i is the length of C .

19. Using (8), find an upper bound for the absolute value of the work W done by the force $\mathbf{p} = xi + y^2j$ in the displacement along the straight line from $(0, 0, 0)$ to $(1, 1, 0)$. Find W by integration and compare the results.
20. Using the trapezoidal rule (Sec. 0.8), evaluate $\int_C f(x, y) ds$ along the straight line $y = x$ in the xy -plane from $(0, 0)$ to $(1, 1)$, where $f(x, y)$ is given by the values $f(0, 0) = 1.0$, $f(\frac{1}{4}, \frac{1}{4}) = 1.5$, $f(\frac{1}{2}, \frac{1}{2}) = 1.7$, $f(\frac{3}{4}, \frac{3}{4}) = 1.5$, $f(1, 1) = 1.0$.

6.3 Double Integrals

In our further consideration we shall need the concept of a double integral. Although the reader will be familiar with double integrals from calculus, we shall now present a brief review.

In the case of a definite integral

$$\int_a^b f(x) dx$$

the integrand is a function $f(x)$ which exists for all x in an interval $a \leq x \leq b$ of the x -axis. In the case of a double integral the integrand will be a function

$f(x, y)$ which is given for all (x, y) in a closed bounded² region R of the xy -plane.

The definition of the double integral is quite similar to that of the definite integral. We subdivide the region R by drawing parallels to the x and y axes (Fig. 175, p. 308). We number those rectangles which are within R from 1 to n . In each such rectangle we choose a point, say, (x_k, y_k) in the k th rectangle, and then we form the sum

$$J_n = \sum_{k=1}^n f(x_k, y_k) \Delta A_k$$

where ΔA_k is the area of the k th rectangle. This we do for larger and larger positive integers n in a completely independent manner but so that the length of the maximum diagonal of the rectangles approaches zero as n approaches infinity. In this fashion we obtain a sequence of real numbers J_{n_1}, J_{n_2}, \dots . Assuming that $f(x, y)$ is continuous in R and R is bounded by finitely many smooth curves (cf. Sec. 6.1) it can be shown³ that the sequence converges and its limit is independent of the choice of subdivisions and corresponding points (x_k, y_k) . This limit is called the **double integral of $f(x, y)$ over the region R** , and is denoted by the symbol

$$\iint_R f(x, y) dx dy.$$

From the definition it follows that double integrals enjoy properties which are quite similar to those of the definite integral. Let f and g be functions of x and y , defined and continuous in a region R . Then

$$\iint_R kf dx dy = k \iint_R f dx dy \quad (k \text{ constant})$$

$$(1) \quad \begin{aligned} \iint_R (f + g) dx dy &= \iint_R f dx dy + \iint_R g dx dy \\ \iint_R f dx dy &= \iint_{R_1} f dx dy + \iint_{R_2} f dx dy \end{aligned} \quad (\text{cf. Fig. 176} \\ \text{on p. 308}).$$

Furthermore, there exists at least one point (x_0, y_0) in R such that

$$(2) \quad \iint_R f(x, y) dx dy = f(x_0, y_0)A,$$

where A is the area of R ; this is called the **mean value theorem for double integrals**.

² "Closed" means that the boundary is part of the region, and "bounded" means that the region can be enclosed in a circle of sufficiently large radius.

³ Cf. Ref. [A1] in Appendix 1.

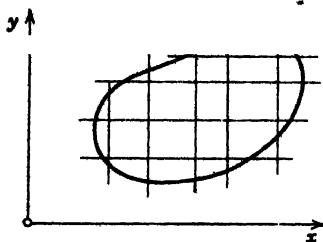
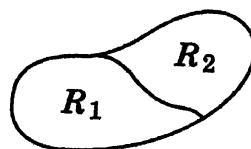
Fig. 175. Subdivision of R .

Fig. 176. Formula (1).

Double integrals over a region R may be evaluated by two successive integrations as follows.

Suppose that R can be described by inequalities of the form

$$a \leq x \leq b, \quad g(x) \leq y \leq h(x)$$

(Fig. 177) so that $y = g(x)$ and $y = h(x)$ represent the boundary of R . Then

$$(3) \quad \iint_R f(x, y) dx dy = \int_a^b \left[\int_{g(x)}^{h(x)} f(x, y) dy \right] dx.$$

We first integrate the inner integral

$$\int_{g(x)}^{h(x)} f(x, y) dy.$$

In this definite integral, x plays the role of a parameter, and the result of the integration will be a function of x , say, $F(x)$. By integrating $F(x)$ over x from a to b we then obtain the value of the double integral in (3).

Similarly, if R can be described by inequalities of the form

$$c \leq y \leq d, \quad p(y) \leq x \leq q(y)$$

(Fig. 178), then we obtain

$$(4) \quad \iint_R f(x, y) dx dy = \int_c^d \left[\int_{p(y)}^{q(y)} f(x, y) dx \right] dy;$$

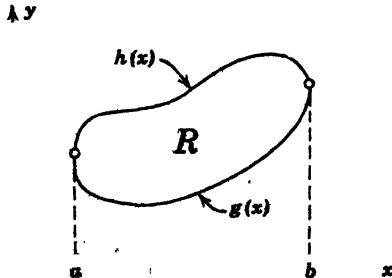


Fig. 177. Evaluation of a double integral.

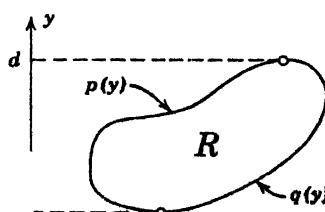


Fig. 178. Evaluation of a double integral.

we now integrate first over x and then the resulting function of y over y from c to d .

If R cannot be represented by those inequalities, then we first subdivide R into suitable portions which have that property, integrate $f(x, y)$ over each portion separately, and finally add the results; this will give us the value of the double integral of $f(x, y)$ over that region R .

Double integrals have various geometrical and physical applications. For example, the area A of R is

$$A = \iint_R dx dy.$$

The volume V beneath the surface $z = f(x, y) (> 0)$ and above the region R in the xy -plane is (Fig. 179)

$$V = \iint_R f(x, y) dx dy,$$

because the term $f(x_k, y_k) \Delta A_k$ in J_n on p. 307 represents the volume of a rectangular parallelepiped with base ΔA_k and altitude $f(x_k, y_k)$.

Let $f(x, y)$ be the density (= mass per unit area) of a distribution of mass in the xy -plane. Then the total mass M in R is

$$M = \iint_R f(x, y) dx dy,$$

the center of gravity of the mass in R has the coordinates

$$\bar{x} = \frac{1}{M} \iint_R xf(x, y) dx dy \quad \text{and} \quad \bar{y} = \frac{1}{M} \iint_R yf(x, y) dx dy,$$

the moments of inertia I_x and I_y of the mass in R about the x and y axes, respectively, are

$$I_x = \iint_R y^2 f(x, y) dx dy, \quad I_y = \iint_R x^2 f(x, y) dx dy,$$

and the polar moment of inertia about the origin is

$$I_0 = I_x + I_y = \iint_R (x^2 + y^2) f(x, y) dx dy.$$

Example 1. Let $f(x, y) = 1$ be the density of mass in the region $R: 0 \leq y \leq \sqrt{1 - x^2}$, $0 \leq x \leq 1$ (Fig. 180). Find the center of gravity and the moments of inertia I_x , I_y , and I_0 .

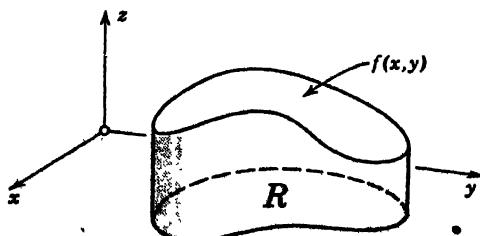


Fig. 179. Double integral as volume.

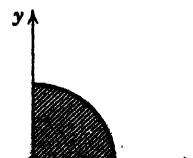


Fig. 180. Example 1.

The total mass in R is

$$M = \iint_R dx dy = \int_0^1 \left[\int_0^{\sqrt{1-x^2}} dy \right] dx = \int_0^1 \sqrt{1-x^2} dx = \int_0^{\pi/2} \cos^2 \theta d\theta = \frac{\pi}{4}$$

($x = \sin \theta$), which is the area of R . The coordinates of the center of gravity are

$$\bar{x} = \frac{4}{\pi} \iint_R x dx dy = \frac{4}{\pi} \int_0^1 \left[\int_0^{\sqrt{1-x^2}} x dy \right] dx = \frac{4}{\pi} \int_0^1 x \sqrt{1-x^2} dx = -\frac{4}{\pi} \int_1^0 z^2 dz = \frac{4}{3\pi}$$

($\sqrt{1-x^2} = z$), and $\bar{y} = \bar{x}$, for reasons of symmetry. Furthermore,

$$I_x = \iint_R y^2 dx dy = \int_0^1 \left[\int_0^{\sqrt{1-x^2}} y^2 dy \right] dx = \frac{1}{3} \int_0^1 (\sqrt{1-x^2})^3 dx \\ = \frac{1}{3} \int_0^{\pi/2} \cos^4 \theta d\theta = \frac{\pi}{16}, \quad I_y = \frac{\pi}{16}, \quad I_0 = I_x + I_y = \frac{\pi}{8} \approx 0.3927.$$

It will often be necessary to change the variables of integration in double integrals. We know from calculus that in the case of a definite integral

$$\int_a^b f(x) dx$$

a new variable of integration u can be introduced by setting

$$x = x(u),$$

where the function $x(u)$ is continuous and has a continuous derivative in some interval $\alpha \leq u \leq \beta$ such that $x(\alpha) = a$, $x(\beta) = b$ [or $x(\alpha) = b$, $x(\beta) = a$] and $x(u)$ varies between a and b when u varies between α and β . Then

$$(5) \quad \int_a^b f(x) dx = \int_{\alpha}^{\beta} f(x(u)) \frac{dx}{du} du.$$

For example, let $f(x) = \sqrt{1-x^2}$, $a = 0$, $b = 1$, and set $x = \sin u$. Then

$$f[x(u)] = \sqrt{1-\sin^2 u} = \cos u, \quad \frac{dx}{du} = \cos u, \quad \alpha = 0, \quad \beta = \frac{\pi}{2},$$

and

$$\int_0^1 \sqrt{1-x^2} dx = \int_0^{\pi/2} \cos^2 u du = \frac{\pi}{4}.$$

In the case of a double integral

$$\iint_R f(x, y) dx dy$$

new variables of integration u, v can be introduced by setting

$$x = x(u, v), \quad y = y(u, v)$$

where the functions $x(u, v)$ and $y(u, v)$ are continuous and have continuous first partial derivatives in some region R^* in the uv -plane so that each point (u_0, v_0) in R^* corresponds to a point $[x(u_0, v_0), y(u_0, v_0)]$ in R and conversely, and furthermore the Jacobian⁴

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

is either positive throughout R^* or negative throughout R^* . Then

$$(6) \quad \iint_R f(x, y) dx dy = \iint_{R^*} f[x(u, v), y(u, v)] \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv;$$

that is, the integrand is expressed in terms of u and v , and $dx dy$ is replaced by the absolute value of the Jacobian J times $du dv$. For the proof see Ref. [A12] in Appendix 1.

For example, polar coordinates r and θ can be introduced by setting

$$x = r \cos \theta, \quad y = r \sin \theta.$$

Then

$$J = \frac{\partial(x, y)}{\partial(r, \theta)} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r,$$

and

$$(7) \quad \iint_R f(x, y) dx dy = \iint_{R^*} f(r \cos \theta, r \sin \theta) r dr d\theta,$$

where R^* is the region in the $r\theta$ -plane corresponding to R in the xy -plane.

Example 2. Using (7), we obtain for I_x in Ex. 1

$$I_x = \iint_R y^2 dx dy = \int_0^{\pi/2} \int_0^1 r^2 \sin^2 \theta r dr d\theta = \int_0^{\pi/2} \sin^2 \theta d\theta \int_0^1 r^3 dr = \frac{\pi}{4} \cdot \frac{1}{4} = \frac{\pi}{16}.$$

Example 3. Evaluate the double integral

$$\iint_R (x^2 + y^2) dx dy$$

where R is the square in Fig. 181. The shape of R suggests the transformation $x + y = u$, $x - y = v$. Then $x = \frac{1}{2}(u + v)$, $y = \frac{1}{2}(u - v)$, the Jacobian is

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2},$$

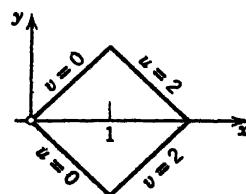


Fig. 181. Region in Example 3.

⁴Named after the German mathematician CARL GUSTAV JACOB JACOBI (1804–1851), who became particularly known by his investigations on elliptic functions.

R corresponds to the square $0 \leq u \leq 2$, $0 \leq v \leq 2$, and, therefore,

$$\iint_R (x^2 + y^2) dx dy = \int_0^2 \int_0^2 \frac{1}{2}(u^2 + v^2) \frac{1}{2} du dv = \frac{8}{3}.$$

Problems

Evaluate and describe the region of integration.

1. $\int_0^1 \int_0^1 (x^2 + y^2) dy dx$

2. $\int_0^1 \int_0^{\sqrt{1-x^2}} (x^2 + y^2) dy dx$

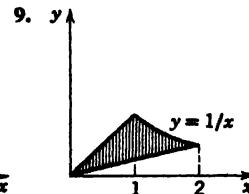
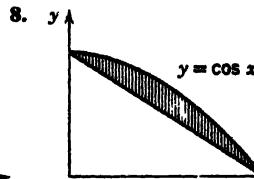
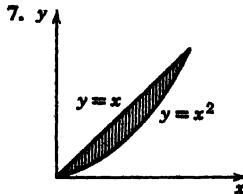
3. $\int_0^1 \int_x^{2x} (x^2 + y^2) dy dx$

4. $\int_{-\pi}^{\pi} \int_{-1}^1 xy dx dy$

5. $\int_0^1 \int_y^{y^2+1} x^2 y dx dy$

6. $\int_0^{\pi} \int_0^{1-\cos \theta} r dr d\theta$

Find the area of the following regions in the xy -plane.



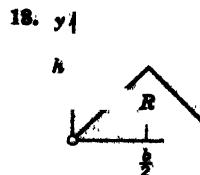
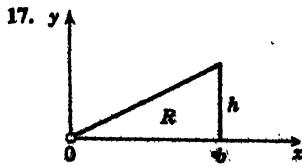
Find the volume of the following regions in space.

10. The tetrahedron cut from the first octant by the plane $x + y + z = 1$
11. The tetrahedron cut from the first octant by the plane $6x + 3y + 2z = 6$
12. The region bounded by the cylinders $x^2 + y^2 = 1$ and $y^2 + z^2 = 1$
13. The region bounded by the surfaces $y = x^2$, $x = y^2$ and the planes $z = 0$, $z = 3$

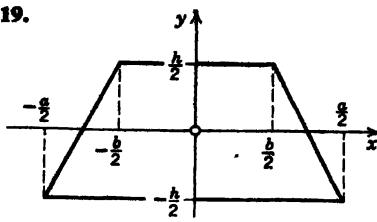
Find the coordinates \bar{x} , \bar{y} of the center of gravity of a mass of density $f(x, y)$ in a region R , where

14. $f(x, y) = 1$, R : the rectangle $0 \leq x \leq 2$, $0 \leq y \leq 4$
15. $f(x, y) = xy$, R : the rectangle $0 \leq x \leq 4$, $0 \leq y \leq 2$
16. $f(x, y) = x^2 + y^2$, R : the region $x^2 + y^2 \leq 1$ in the first quadrant

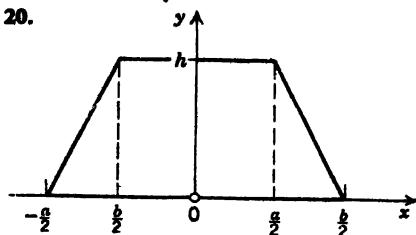
Find the moments of inertia I_x , I_y , I_0 of a mass of density $f(x, y) = 1$ in a region R shown in the following figures.



19.



20.



Using polar coordinates, evaluate $\iint_R f(x, y) dx dy$, where

$$21. f = x + y, \quad R: x^2 + y^2 \leq 9, \quad x \geq 0$$

$$22. f = \cos(x^2 + y^2), \quad R: x^2 + y^2 \leq \pi/2, \quad x \geq 0$$

$$23. f = (x^2 + y^2)^2, \quad R: x^2 + y^2 \leq 4, \quad y \geq 0$$

$$24. f = e^{-x^2-y^2}, \quad R: x^2 + y^2 \leq 1$$

$$25. f = e^{-x^2-y^2}, \quad R: \text{the annulus bounded by } x^2 + y^2 = 1 \text{ and } x^2 + y^2 = 4$$

6.4 Transformation of Double Integrals into Line Integrals

Double integrals over a plane region may be transformed into line integrals over the boundary of the region and conversely. This transformation is of practical as well as theoretical interest and can be done by means of the following basic theorem.

Green's⁵ theorem in the plane. Let R be a closed bounded region in the xy -plane whose boundary C consists of finitely many smooth curves. Let $f(x, y)$ and $g(x, y)$ be functions which are continuous and have continuous partial derivatives $\partial f / \partial y$ and $\partial g / \partial x$ everywhere in some domain containing R . Then

$$(1) \quad \iint_R \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) dx dy = \int_C (f dx + g dy),$$

the integration being taken along the entire boundary C of R such that R is on the left as one advances in the direction of integration (cf. Fig. 182).

Proof. We first prove the theorem for a special

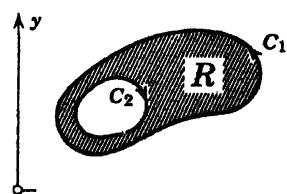


Fig. 182. Region R whose boundary C consists of two parts; C_1 is traversed in the counterclockwise sense, while C_2 is traversed in the clockwise sense.

*GEORGE GREEN (1793–1841), English mathematician.

region R which can be represented in both of the forms

$$a \leq x \leq b, \quad u(x) \leq y \leq v(x),$$

and $c \leq y \leq d, \quad p(y) \leq x \leq q(y)$

(Fig. 183). Using (3) in Sec. 6.3, we obtain

$$(2) \quad \iint_R \frac{\partial f}{\partial y} dx dy = \int_a^b \left[\int_{u(x)}^{v(x)} \frac{\partial f}{\partial y} dy \right] dx.$$

We integrate the inner integral:

$$\int_{u(x)}^{v(x)} \frac{\partial f}{\partial y} dy = f(x, y) \Big|_{y=u(x)}^{y=v(x)} = f[x, v(x)] - f[x, u(x)].$$

By inserting this into (2) we find

$$\begin{aligned} \iint_R \frac{\partial f}{\partial y} dx dy &= \int_a^b f[x, v(x)] dx - \int_a^b f[x, u(x)] dx \\ &= - \int_a^b f[x, u(x)] dx - \int_b^a f[x, v(x)] dx. \end{aligned}$$

Since $y = u(x)$ represents the oriented curve C^* (Fig. 183a) and $y = v(x)$ represents C^{**} , the integrals on the right may be written as line integrals over C^* and C^{**} , and, therefore,

$$(3) \quad \begin{aligned} \iint_R \frac{\partial f}{\partial y} dx dy &= - \int_{C^*} f(x, y) dx - \int_{C^{**}} f(x, y) dx \\ &= - \int_C f(x, y) dx. \end{aligned}$$

If portions of C are segments parallel to the y -axis (such as C and \bar{C} in Fig. 184), then the result is the same as before, because the integrals over these portions are zero and may be added to the integrals over C^* and C^{**} .

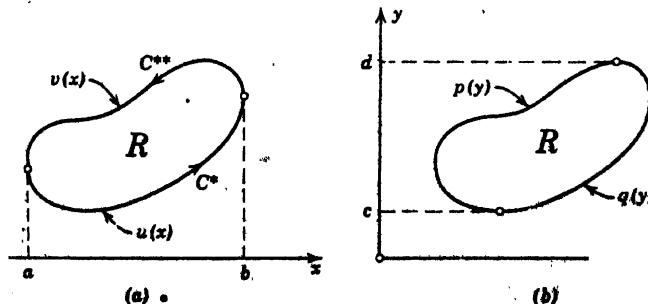


Fig. 183. Example of a special region for Green's theorem.

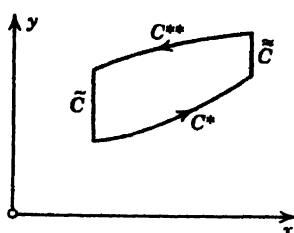


Fig. 184. Proof of Green's theorem.

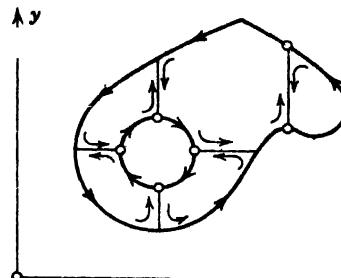


Fig. 185. Proof of Green's theorem.

to obtain the integral over the whole boundary C in (3). Similarly, using (4), Sec. 6.3, we obtain (cf. Fig. 183b)

$$\iint_R \frac{\partial g}{\partial x} dx dy = \int_C \left[\int_{\nu(y)}^{u(y)} \frac{\partial g}{\partial x} dx \right] dy = \int_C g(x, y) dy.$$

From this and (3), the formula (1) follows, and the theorem is proved for special regions.

We now prove the theorem for a region R which itself is not a special region but can be subdivided into finitely many special regions (Fig. 185). In this case we apply the theorem to each subregion and then add the results; the left-hand members add up to the integral over R while the right-hand members add up to the line integral over C plus integrals over the curves introduced for subdividing R . Each of the latter integrals occurs twice, taken once in each direction. Hence these two integrals cancel each other, and we are left with the line integral over C .

The proof thus far covers all regions which are of interest in engineering problems. To prove the theorem for the most general region R satisfying the conditions in the theorem, we must approximate R by a region of the type just considered and then use a limiting process. For details see Ref. [A12] in Appendix 1.

Green's theorem will be of basic importance in our further consideration. For the time being we consider a few simple illustrative examples.

Example 1 (Area of a plane region as a line integral over the boundary). In (1), let $f = 0$ and $g = x$. Then

$$\iint_R dx dy = \int_C x dy.$$

The integral on the left is the area A of R . Similarly, let $f = -y$ and $g = 0$; then from (1),

$$= \iint_R dx dy = - \int_C y dx.$$

By adding both formulas we obtain

$$(4) \quad \frac{1}{2} \int_C (x \, dy - y \, dx),$$

the integration being taken as indicated in Green's theorem. This interesting formula expresses the area of R in terms of a line integral over the boundary. It has various applications; for example, the theory of certain planimeters (cf. Sec. 0.8) is based upon this formula.

Example 2 (Area of a plane region in polar coordinates). Let r and θ be polar coordinates defined by $x = r \cos \theta$, $y = r \sin \theta$. Then

$$dx = \cos \theta \, dr - r \sin \theta \, d\theta, \quad dy = \sin \theta \, dr + r \cos \theta \, d\theta,$$

and (4) assumes the form

$$(5) \quad A = \frac{1}{2} \int_C r^2 \, d\theta.$$

This formula is well known from calculus.

As an application of (5), consider the cardioid $r = a(1 - \cos \theta)$ where $0 \leq \theta \leq 2\pi$ (Fig. 186). We find

$$A = \frac{a^2}{2} \int_0^{2\pi} (1 - \cos \theta)^2 \, d\theta = \frac{3\pi}{2} a^2.$$

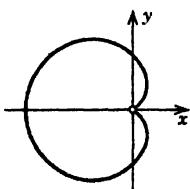


Fig. 186. Cardioid.

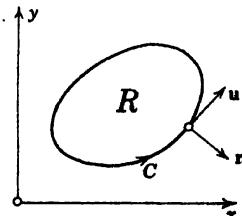


Fig. 187. Example 3.

✓ Example 3 (Transformation of a double integral of the Laplacian of a function into a line integral of its normal derivative). Let $w(x, y)$ be a function which is continuous and has continuous first and second derivatives in some region R of the xy -plane where R is of the type indicated in Green's theorem. We set $f = -\partial w / \partial y$ and $g = \partial w / \partial x$. Then $\partial f / \partial y$ and $\partial g / \partial x$ are continuous in R , and

$$(6) \quad \frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} = \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = \nabla^2 w,$$

the Laplacian of w (cf. Sec. 5.16). Furthermore,

$$(7) \quad \int_C (f \, dx + g \, dy) = \int_C \left(f \frac{dx}{ds} + g \frac{dy}{ds} \right) ds = \int_C \left(-\frac{\partial w}{\partial y} \frac{dx}{ds} + \frac{\partial w}{\partial x} \frac{dy}{ds} \right) ds,$$

where s is the arc length of C , and C is oriented as shown in Fig. 187. The integrand of the last integral may be written as the scalar product of the vectors

$$\text{grad } w = \frac{\partial w}{\partial x} \mathbf{i} + \frac{\partial w}{\partial y} \mathbf{j} \quad \text{and} \quad \mathbf{n} = \frac{dy}{ds} \mathbf{i} - \frac{dx}{ds} \mathbf{j};$$

that is,

$$(8) \quad -\frac{\partial w}{\partial y} \frac{dx}{ds} + \frac{\partial w}{\partial x} \frac{dy}{ds} = (\text{grad } w) \cdot \mathbf{n}.$$

The vector \mathbf{n} is a unit normal to C , because the vector

$$\mathbf{u} = \frac{dr}{ds} = \frac{dx}{ds} \mathbf{i} + \frac{dy}{ds} \mathbf{j}$$

(cf. Sec. 5.13) is the unit tangent vector to C , and $\mathbf{u} \cdot \mathbf{n} = 0$. Furthermore, it is not difficult to see that \mathbf{n} is directed to the exterior of C . From this and (5) in Sec. 5.16 it follows that the expression on the right-hand side of (8) is the derivative of w in the direction of the outward normal to C . Denoting this directional derivative by $\partial w / \partial n$ and taking (6), (7), and (8) into account, we obtain from Green's theorem the useful integral formula

$$(9) \quad \iint_R \nabla^2 w \, dx \, dy = \int_C \frac{\partial w}{\partial n} \, ds.$$

Further important applications and consequences of Green's theorem in the plane will be considered in the following sections.

Problems

Using (a) direct calculation, (b) Green's theorem, evaluate

1. $\int_C (y^2 dx + x^2 dy)$, C : the boundary of the square $-1 \leq x \leq 1$, $-1 \leq y \leq 1$ (counterclockwise)
2. $\int_C (y \, dx - x \, dy)$, C : the boundary of the square $0 \leq x \leq 1$, $0 \leq y \leq 1$ (clockwise)
3. $\int_C [(3x^2 + y) \, dx + 4y^2 \, dy]$, C : the boundary of the triangle with vertices $(0, 0)$, $(1, 0)$, $(0, 2)$ (counterclockwise)

Using Green's theorem, evaluate the line integral $\int_C (f \, dx + g \, dy)$ counterclockwise around the given contour C , where $f \, dx + g \, dy$ equals

4. $(x - y) \, dx - x^2 \, dy$, C : the boundary of the square $0 \leq x \leq 2$, $0 \leq y \leq 2$
5. $(x^3 - 3y) \, dx + (x + \sin y) \, dy$, C as in Prob. 3
6. $(e^x - 3y) \, dx + (e^y + 6x) \, dy$, C : $x^2 + 4y^2 = 4$
7. $(x^2 - \cosh y) \, dx + (y + \sin x) \, dy$, C : the boundary of the rectangle $0 \leq x \leq \pi$, $0 \leq y \leq 1$
8. $-y^3 \, dx + x^3 \, dy$, C : $x^2 + y^2 = 1$
9. $(\cos x \sin y - xy) \, dx + \sin x \cos y \, dy$, C as in Prob. 8
10. $2 \arctan(y/x) \, dx + \ln(x^2 + y^2) \, dy$, C : $(x - 1)^2 + (y + 1)^2 = 4$
11. $yx^2 \, dx + x^2 \, dy$, C : the boundary of the triangle with vertices $(0, 0)$, $(0, 1)$, $(1, 1)$
12. $e^x \sin y \, dx + e^y \cos y \, dy$, C : $9x^2 + y^2 = 9$
13. $x^{-1}e^y \, dx + (e^y \ln x + 2x) \, dy$, C : the boundary of the region bounded by $y = x^4 + 1$ and $y = 2$

Using one of the formulas in Ex. 1, find the area of the following plane regions.

14. The interior of the ellipse $x^2/a^2 + y^2/b^2 = 1$

15. The region in the first quadrant bounded by $y = x$, $y = 1/x$, and $y = x/4$.
 16. The region bounded by $y = x^2$ and $y = x + 2$

Using (9), evaluate $\int_C \frac{\partial w}{\partial n} ds$ where

$$17. w = x^2 + 3y^2, \quad C: x^2 + y^2 = 4$$

$$18. w = 3x^2y - y^3, \quad C \text{ as in Prob. 2}$$

$$19. w = e^x + e^y, \quad C: \text{the boundary of the rectangle } 0 \leq x \leq 2, 0 \leq y \leq 1$$

20. If $w(x, y)$ satisfies Laplace's equation $\nabla^2 w = 0$ in a region R , show that

$$(10) \quad \iint_R \left[\left(\frac{\partial w}{\partial x} \right)^2 + \left(\frac{\partial w}{\partial y} \right)^2 \right] dx dy = \int_C w \frac{\partial w}{\partial n} ds$$

where $\partial w / \partial n$ is defined as in Ex. 3. Hint: model your work somewhat after that of Ex. 3.

Using (10), evaluate $\int_C w \frac{\partial w}{\partial n} ds$ where

$$21. w = x, \quad C \text{ as in Prob. 2}$$

$$22. w = e^x \cos y, \quad C \text{ as in Prob. 2}$$

23. Check the results of Probs. 21 and 22 by direct calculation.

24. Introducing $\mathbf{v} = g\mathbf{i} - f\mathbf{j}$, show that the formula in Green's theorem may be written

$$(11) \quad \iint_R \operatorname{div} \mathbf{v} dx dy = \int_C \mathbf{v} \cdot \mathbf{n} ds$$

where \mathbf{n} is the outward unit normal vector to C (Fig. 187) and s is the arc length of C .

25. Show that the formula in Green's theorem may be written

$$(12) \quad \iint_R (\operatorname{curl} \mathbf{v}) \cdot \mathbf{k} dx dy = \int_C \mathbf{v} \cdot \mathbf{u} ds$$

where \mathbf{k} is a unit vector perpendicular to the xy -plane, \mathbf{u} is the unit tangent vector to C , and s is the arc length of C .

6.5 Surfaces

In Sec. 6.7 we shall consider surface integrals. This consideration will require knowledge of some basic facts about surfaces which we shall now explain and illustrate by simple examples.

A surface S may be represented in the form

$$(1) \quad f(x, y, z) = 0,$$

where x, y, z are Cartesian coordinates in space. Then the gradient of f is normal to S (cf. Theorem 2 in Sec. 5.16), provided $\operatorname{grad} f \neq 0$. Consequently, for S to have a unique normal at each point whose direction depends continuously on the points of S , we must require that f has continuous first

partial derivatives, and at each point at least one of these three derivatives is not zero. Then the vector

$$(2) \quad \mathbf{n} = \frac{\operatorname{grad} f}{|\operatorname{grad} f|}$$

is a unit normal vector of S (and $-\mathbf{n}$ is the other).

Example 1. The sphere $x^2 + y^2 + z^2 - a^2 = 0$ has the unit normal vector

$$\mathbf{n}(x, y, z) = \frac{x}{a} \mathbf{i} + \frac{y}{a} \mathbf{j} + \frac{z}{a} \mathbf{k}.$$

Sometimes it is convenient to use an explicit representation

$$(3) \quad z = g(x, y)$$

of a given surface. Clearly, by writing $z - g(x, y) = 0$ we immediately obtain an implicit representation of the form (1).

In Sec. 5.11 the usefulness of parametric representations of curves was demonstrated. We shall now see that surfaces may also be represented in parametric form. Suppose that

$$(4) \quad \mathbf{r}(u, v) = x(u, v)\mathbf{i} + y(u, v)\mathbf{j} + z(u, v)\mathbf{k} \quad \checkmark$$

is a vector function of two independent real variables u and v which is single-valued and continuous in a simply-connected bounded region R of the uv -plane and is at least once continuously differentiable in R . Then to any point (u_0, v_0) in R there corresponds a point with position vector $\mathbf{r}(u_0, v_0)$ in space. Furthermore, we assume that

$$(5) \quad \mathbf{r}_u \times \mathbf{r}_v \neq \mathbf{0}$$

everywhere in R ; here, and in the following, subscripts u and v denote partial derivatives. We shall see that a vector function $\mathbf{r}(u, v)$ satisfying these conditions represents a surface S in space which has a unique normal whose direction depends continuously on the points of S . The representation (4) is called a **parametric representation** of S , and the variables u and v are called the **parameters** of this representation or the **coordinates on S** . The curves $u = \text{const}$ and $v = \text{const}$ form a net on S ; they are called **coordinate curves on S** .

If a surface S has a unique normal whose direction depends continuously on the points of S , then S is called a **smooth surface**.

If S is not smooth but can be subdivided into finitely many smooth portions, then it is called a **piecewise smooth surface**.

For example, the surface of a sphere is smooth, while the surface of a cube is piecewise smooth.

Example 2. The sphere in Ex. 1 may be represented in the form

$$(6) \quad \mathbf{r}(u, v) = a \cos v \cos u \mathbf{i} + a \cos v \sin u \mathbf{j} + a \sin v \mathbf{k}$$

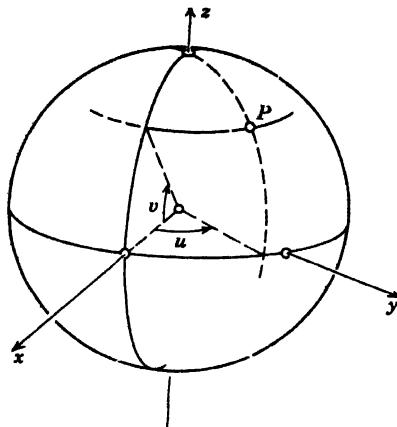


Fig. 188. Parametric representation of the sphere.

where $0 \leq u \leq 2\pi$, $-\pi/2 \leq v \leq \pi/2$; that is,

$$x = a \cos v \cos u, \quad y = a \cos v \sin u, \quad z = a \sin v.$$

The curves $u = \text{const}$ and $v = \text{const}$ are the "meridians" and "parallels" on S (cf. Fig. 188). The relation (5) is satisfied everywhere except at the "poles" $v = -\pi/2$ and $v = \pi/2$. The representation (6) is used in geography for measuring the latitude and longitude of points on the globe.

Problems

What surfaces are represented by the following parametric representations? What are the coordinate curves?

- | | |
|--|---|
| 1. $\mathbf{r} = u\mathbf{i} + v\mathbf{j}$ | 2. $\mathbf{r} = u \cos v \mathbf{i} + u \sin v \mathbf{j}$ |
| 3. $\mathbf{r} = \cos u \mathbf{i} + \sin u \mathbf{j} + vk$ | 4. $\mathbf{r} = u\mathbf{i} + v\mathbf{j} + uk$ |
| 5. $\mathbf{r} = 4 \cos u \mathbf{i} + \sin u \mathbf{j} + vk$ | 6. $\mathbf{r} = u\mathbf{i} + v\mathbf{j} + (u + v)\mathbf{k}$ |
| 7. $\mathbf{r} = u \cos v \mathbf{i} + u \sin v \mathbf{j} + uk$ | 8. $\mathbf{r} = u \cos v \mathbf{i} + u \sin v \mathbf{j} + u^2\mathbf{k}$ |

Find a parametric representation of the following surfaces.

- | | |
|--------------------------------------|--|
| 9. The xz -plane | 10. The plane $z = x$ |
| 11. The plane $x + y + z = 1$ | 12. The cylinder of revolution $y^2 + z^2 = 1$ |
| 13. The parabolic cylinder $z = x^2$ | 14. The elliptic cylinder $y^2 + 9z^2 = 9$ |

Represent the following surfaces in the form (1).

- | |
|---|
| 15. The ellipsoid $\mathbf{r} = a \cos v \cos u \mathbf{i} + b \cos v \sin u \mathbf{j} + c \sin v \mathbf{k}$ |
| 16. The elliptic paraboloid $\mathbf{r} = au \cos v \mathbf{i} + bu \sin v \mathbf{j} + u^2\mathbf{k}$ |
| 17. The hyperbolic paraboloid $\mathbf{r} = au \cosh v \mathbf{i} + bu \sinh v \mathbf{j} + u^2\mathbf{k}$ |
| 18. The hyperboloid $\mathbf{r} = a \sinh u \cos v \mathbf{i} + b \sinh u \sin v \mathbf{j} + c \cosh u \mathbf{k}$ |
| 19. Sketch figures of the surfaces in Probs. 15–18. |
| 20. Find the points in Ex. 2, Prob. 2, and Prob. 7 at which (5) is not satisfied. |

6.6 Tangent Plane. First Fundamental Form. Area

If a surface S is represented in the form $\mathbf{r} = \mathbf{r}(u, v)$, then a curve on S can be represented by a pair of continuous functions

$$(1) \quad u = g(t), \quad v = h(t)$$

of a real parameter t .

Example 1. The vector function $\mathbf{r}(u, v) = a \cos u \mathbf{i} + a \sin u \mathbf{j} + v\mathbf{k}$ represents a cylinder of revolution S of radius a . The equations $u = t$, $v = ct$ represent a circular helix on S . In fact, by substituting these equations into the representation of S we obtain (cf. Ex. 3 in Sec. 5.11)

$$\mathbf{r}[u(t), v(t)] = a \cos t \mathbf{i} + a \sin t \mathbf{j} + ct\mathbf{k}.$$

Let S be a smooth surface represented by a vector function $\mathbf{r}(u, v)$ and let C be a curve on S represented in the form (1). Then C , considered as a curve in space, is represented by the vector function

$$(2) \quad \mathbf{r}(t) = \mathbf{r}[u(t), v(t)].$$

Suppose that both functions in (1) have continuous first derivatives such that for each t at least one of these derivatives is not zero. Then C has a tangent at each of its points whose direction depends continuously on the points, and a tangent vector to C is

$$\dot{\mathbf{r}}(t) \equiv \frac{d\mathbf{r}}{dt} = \mathbf{r}_u \dot{u} + \mathbf{r}_v \dot{v}.$$

From (5) in the last section it follows that the vectors $\mathbf{r}_u \equiv \partial \mathbf{r} / \partial u$ and $\mathbf{r}_v \equiv \partial \mathbf{r} / \partial v$ are linearly independent. Hence they determine a plane. This plane is called the tangent plane of S at the corresponding point P of S and

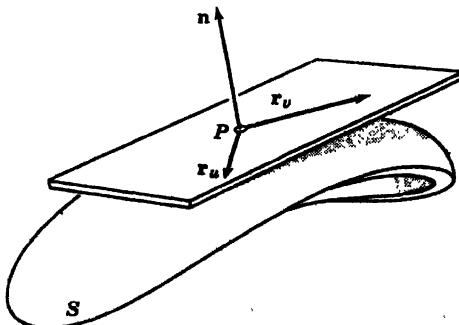


Fig. 189. Tangent plane and normal vector.

will be denoted by $T(P)$. It touches S at P . From (2) it follows that $T(P)$ contains the tangent to any curve on S through P , at that point P (see also Sec. 5.16).

The straight line through P and perpendicular to $T(P)$ is called the **normal** to S at P . Since the vectors \mathbf{r}_u and \mathbf{r}_v lie in $T(P)$, the unit vector

$$(3) \quad \mathbf{n} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|}$$

(Fig. 189) is perpendicular to $T(P)$. This vector is called the **unit normal vector** of S at P . Its sense depends on the choice of the coordinates u and v ; the transformation $u = -\bar{u}$, $v = \bar{v}$ or any other transformation whose Jacobian (cf. Sec. 6.3) has a negative value, reverses the sense of \mathbf{n} .

We shall now determine the *linear element* of a curve C , represented in the form (1), on a surface S , represented by $\mathbf{r}(u, v)$. We have

$$d\mathbf{r} = \mathbf{r}_u du + \mathbf{r}_v dv,$$

and, by (4) in Sec. 5.12,

$$\begin{aligned} ds^2 &= d\mathbf{r} \cdot d\mathbf{r} = (\mathbf{r}_u du + \mathbf{r}_v dv) \cdot (\mathbf{r}_u du + \mathbf{r}_v dv) \\ &= \mathbf{r}_u \cdot \mathbf{r}_u du^2 + 2 \mathbf{r}_u \cdot \mathbf{r}_v du dv + \mathbf{r}_v \cdot \mathbf{r}_v dv^2. \end{aligned}$$

If we use the standard notation

$$(4) \quad E = \mathbf{r}_u \cdot \mathbf{r}_u, \quad F = \mathbf{r}_u \cdot \mathbf{r}_v, \quad G = \mathbf{r}_v \cdot \mathbf{r}_v,$$

the last expression assumes the form

$$(5) \quad ds^2 = E du^2 + 2F du dv + G dv^2.$$

This quadratic differential form is called the **first fundamental form** of S .

Example 2. The vector function

$$\mathbf{r}(u, v) = u \cos v \mathbf{i} + u \sin v \mathbf{j}$$

represents the xy -plane, and u and v are polar coordinates. We find

$$\mathbf{r}_u = \cos v \mathbf{i} + \sin v \mathbf{j}, \quad \mathbf{r}_v = -u \sin v \mathbf{i} + u \cos v \mathbf{j}.$$

Hence $E = 1$, $F = 0$, $G = u^2$, and the first fundamental form corresponding to polar coordinates $u = \rho$ and $v = \theta$ is

$$(6) \quad ds^2 = d\rho^2 + \rho^2 d\theta^2.$$

We shall now see that the first fundamental form is of basic importance because it enables us to measure lengths, angles between curves, and areas on the corresponding surface S .

From (1) and (4) in Sec. 5.12 and the preceding formula (5) it follows that a curve

$$C: \quad u(t), v(t), \quad a \leq t \leq b,$$

on a surface S : $\mathbf{r}(u, v)$ has the length

$$(7) \quad l = \int_a^b \sqrt{\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}} dt = \int_a^b \frac{ds}{dt} dt = \int_a^b \sqrt{E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2} dt.$$

We now consider two curves

$$C_1: u = g(t), v = h(t) \quad \text{and} \quad C_2: u = p(t), v = q(t)$$

on a surface S : $\mathbf{r}(u, v)$ which intersect at a point P of S . The vectors

$$\mathbf{a} = \frac{d}{dt} \mathbf{r}[g(t), h(t)] = \mathbf{r}_u \dot{g} + \mathbf{r}_v \dot{h}$$

and

$$\mathbf{b} = \frac{d}{dt} \mathbf{r}[p(t), q(t)] = \mathbf{r}_u \dot{p} + \mathbf{r}_v \dot{q}$$

at P are tangent to C_1 and C_2 , respectively. The angle of intersection between C_1 and C_2 at P is defined as the angle γ between \mathbf{a} and \mathbf{b} , and from (3) in Sec. 5.4 we have

$$(8) \quad \cos \gamma = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| |\mathbf{b}|}$$

where

$$\mathbf{a} \cdot \mathbf{b} = (\mathbf{r}_u \dot{g} + \mathbf{r}_v \dot{h}) \cdot (\mathbf{r}_u \dot{p} + \mathbf{r}_v \dot{q}) = E\dot{g}\dot{p} + F(\dot{g}\dot{q} + \dot{h}\dot{p}) + G\dot{h}\dot{q}$$

and similarly

$$|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{E\dot{g}^2 + 2F\dot{g}\dot{h} + G\dot{h}^2},$$

$$|\mathbf{b}| = \sqrt{\mathbf{b} \cdot \mathbf{b}} = \sqrt{E\dot{p}^2 + 2F\dot{p}\dot{q} + G\dot{q}^2}.$$

This important result shows that the angle between two intersecting curves on a surface can be expressed in terms of E , F , G and the derivatives of the functions representing the curves, evaluated at the point of intersection.

We have shown that the coefficients E , F , and G of the first fundamental form permit the measurement of lengths and angles on a surface. We shall now demonstrate that they also permit the measurement of areas on a surface.

The area A of a surface S : $\mathbf{r}(u, v)$ is defined by the double integral

$$(9) \quad A = \iint_R \sqrt{EG - F^2} du dv \quad \checkmark$$

over the region R in the uv -plane corresponding to that surface; here the positive sign of the square roots is taken. The expression

$$(10) \quad dA = \sqrt{EG - F^2} du dv$$

is called the element of area of S .

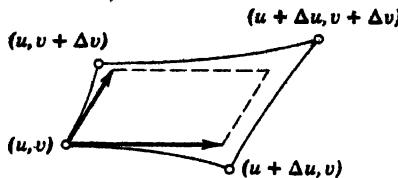


Fig. 190. Area.

The formula (10) can be made plausible by noting that

$$\Delta A = \sqrt{EG - F^2} \Delta u \Delta v$$

is the area of a small parallelogram whose sides are the vectors

$$\mathbf{r}_u \Delta u \quad \text{and} \quad \mathbf{r}_v \Delta v$$

(Fig. 190). Indeed, from the definition of a vector product it follows that

$$\Delta A = |\mathbf{r}_u \Delta u \times \mathbf{r}_v \Delta v| = |\mathbf{r}_u \times \mathbf{r}_v| \Delta u \Delta v,$$

and from (5) in Sec. 5.5 and (4) in the current section we obtain

$$|\mathbf{r}_u \times \mathbf{r}_v|^2 = (\mathbf{r}_u \cdot \mathbf{r}_u)(\mathbf{r}_v \cdot \mathbf{r}_v) - (\mathbf{r}_u \cdot \mathbf{r}_v)^2;$$

that is,

$$(11) \quad |\mathbf{r}_u \times \mathbf{r}_v|^2 = EG - F^2.$$

The integral (9) is obtained by subdividing S into parts S_1, \dots, S_n , approximating each part S_k by a portion of the tangent plane of S at a point in S_k , and forming the sum of all the approximating areas. This is done for each $n = 1, 2, \dots$ such that the dimensions of the largest S_k approach zero as n approaches infinity. The limit of those sums is the integral (9). For details see Ref. [C4] in Appendix 1.

Example 3 (Torus). The vector function

$$\mathbf{r}(u, v) = (a + b \cos v) \cos u \mathbf{i} + (a + b \cos v) \sin u \mathbf{j} + b \sin v \mathbf{k} \quad (a > b > 0)$$

represents a torus (Fig. 191) which is obtained by rotating a circle C about a fixed straight line A so that the plane of C always passes through A , and C does not intersect A . Using (4), we find

$$E = (a + b \cos v)^2, \quad F = 0, \quad G = b^2,$$

$$EG - F^2 = b^2(a + b \cos v)^2,$$

and the total area is

$$A = \int_0^{2\pi} \int_0^{2\pi} b(a + b \cos v) du dv = 4\pi^2 ab.$$

If S is represented in the form

$$z = g(x, y),$$

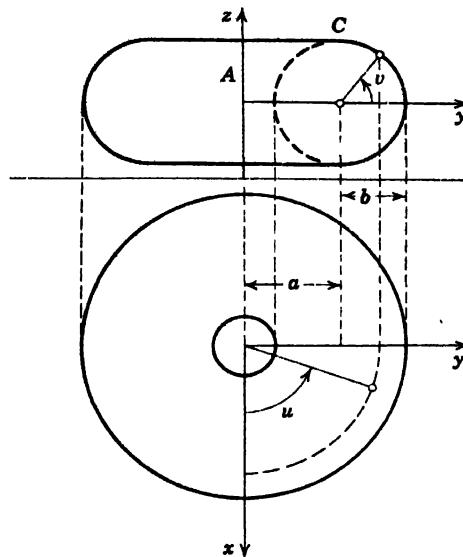


Fig. 191. Torus.

we may set $x = u$ and $y = v$ and write the representation in parametric form:

$$\mathbf{r}(u, v) = u\mathbf{i} + v\mathbf{j} + g(u, v)\mathbf{k}.$$

Then the partial derivatives of \mathbf{r} with respect to u and v are

$$(12) \quad \mathbf{r}_u = \mathbf{i} + g_u \mathbf{k}, \quad \mathbf{r}_v = \mathbf{j} + g_v \mathbf{k}.$$

It follows that the coefficients of the first fundamental form are

$$E = 1 + g_u^2, \quad F = g_u g_v, \quad G = 1 + g_v^2.$$

From this we obtain

$$EG - F^2 = 1 + g_u^2 + g_v^2.$$

Since $u = x$ and $v = y$, the integral (9) now assumes the form

$$(9') \quad A = \iint_S \sqrt{1 + \left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2} dx dy$$

where \bar{S} is the orthogonal projection of S into the xy -plane. Clearly,

$$(10') \quad dA = \sqrt{1 + \left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2} dx dy.$$

We want to show that this may be written

$$(10'') \quad dA = \sec \gamma dx dy$$

where γ is the acute angle between the z -axis and the undirected normal to S .

In fact, the vector product of the vectors in (12) is perpendicular to S ; denoting it by \mathbf{a} and observing that $u = x$, $v = y$, we have

$$\mathbf{a} = \frac{\partial \mathbf{r}}{\partial x} \times \frac{\partial \mathbf{r}}{\partial y} = -\frac{\partial g}{\partial x} \mathbf{i} - \frac{\partial g}{\partial y} \mathbf{j} + \mathbf{k}, \quad |\mathbf{a}| = \sqrt{1 + \left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2}.$$

Furthermore, $\mathbf{a} \cdot \mathbf{k} = |\mathbf{a}| \cos \gamma^* = 1$, where γ^* is the angle between \mathbf{a} and the positive z -axis. Clearly $\cos \gamma^* > 0$ and therefore $\gamma^* < \pi/2$, that is, γ^* is acute and thus equal to γ . Hence

$$|\mathbf{a}| \cos \gamma = 1 \quad \text{or} \quad \sec \gamma = |\mathbf{a}| \quad \left(\gamma < \frac{\pi}{2} \right)$$

and (10") is established.

Problems

Find a representation of the tangent plane T of a surface S at a point P , if S is represented in the form

1. $\mathbf{r} = \mathbf{r}(u, v)$ 2. $z = g(x, y)$ 3. $f(x, y, z) = 0$
 4. Under what conditions do the coordinate curves $p = \text{const}$ and $q = \text{const}$ in the solution of Prob. 1 intersect at right angles?

Find a representation of the tangent plane of the following surfaces at the given point P_0 : (x_0, y_0, z_0) , where

5. $z = xy$, $P_0: (1, 1, 1)$ 6. $x^2 + y^2 + z^2 = 25$, $P_0: (3, 4, 0)$
 7. $z = x^2$, $P_0: (2, 1, 4)$ 8. $x^2 + y^2 = 8$, $P_0: (2, 2, 3)$

Find a representation of the unit normal vector of a surface S represented in the form

9. $z = g(x, y)$ 10. $f(x, y, z) = 0$

11. In (6), Sec. 6.5, let $u = -\dot{u}$, $v = \dot{v}$. Show that the unit normal vector (3) corresponding to the resulting representation is opposite to that corresponding to the original representation (6).

Find the first fundamental form of the following surfaces.

12. $\mathbf{r} = u\mathbf{i} + v\mathbf{j}$ 13. $\mathbf{r} = 2u\mathbf{i} + 3v\mathbf{j}$
 14. $\mathbf{r} = (u + v)\mathbf{i} + (u - v)\mathbf{j}$ 15. $\mathbf{r} = u\mathbf{i} + v\mathbf{j} + u^2\mathbf{k}$
 16. $\mathbf{r} = \cos u \mathbf{i} + \sin u \mathbf{j} + v\mathbf{k}$ 17. $\mathbf{r} = u\mathbf{i} + v\mathbf{j} + uv\mathbf{k}$
 18. $\mathbf{r} = 2 \cos u \mathbf{i} + 2 \sin u \mathbf{j} + v\mathbf{k}$ 19. $\mathbf{r} = u \cos v \mathbf{i} + u \sin v \mathbf{j} + u^2\mathbf{k}$

20. Show that the coordinate curves $u = \text{const}$ and $v = \text{const}$ on a surface $\mathbf{r} = \mathbf{r}(u, v)$ intersect at right angles if, and only if, $F = \mathbf{r}_u \cdot \mathbf{r}_v = 0$. Give examples.

Using (9), find the area of the following surfaces.

21. $x^2 + y^2 = 1$, $0 \leq z \leq 1$ 22. The sphere (6) in Sec. 6.5
 23. $z = x^2 + y^2$, $0 \leq z \leq 1$ 24. $z^2 = x^2 + y^2$, $-1 \leq z \leq 1$
 25. Show that the area A in Ex. 3 can be obtained by the theorem of Pappus, which states that the area of a surface of revolution equals the product of the length of a meridian C and the length of the path of the center of gravity of C when C is rotated through the angle 2π .

6.7 Surface Integrals

The concept of a surface integral is a natural generalization of the concept of a double integral considered in Sec. 6.3. Surface integrals occur in many applications, for example, in connection with the center of gravity of a curved lamina, the potential due to charges distributed on surfaces, etc.

The definition of a surface integral parallels that of a double integral. Let S be a portion of a surface of finite area, and let $f(P) = f(x, y, z)$ be a single-valued function which is defined and continuous on S . We subdivide S into n parts S_1, \dots, S_n of areas $\Delta A_1, \dots, \Delta A_n$. In each part S_k we choose an arbitrary point P_k and form the sum

$$(1) \quad J_n = \sum_{k=1}^n f(P_k) \Delta A_k.$$

This we do for each $n = 1, 2, \dots$ in an arbitrary manner, but so that the largest part S_k shrinks to a point as n approaches infinity. The infinite sequence J_1, J_2, \dots has a limit which is independent of the choice of subdivisions and points P_k ; the proof is similar to that in the case of a double integral. This limit is called the **surface integral** of $f(x, y, z)$ over S and is denoted by

$$(2) \quad \iint_S f(x, y, z) dA.$$

To evaluate the surface integral (2), we may reduce it to a double integral as follows. If S is represented in parametric form by a vector function $\mathbf{r}(u, v)$, then $dA = \sqrt{EG - F^2} du dv$, cf. (10) in the previous section, and

$$(3) \quad \iint_S f(x, y, z) dA = \iint_R f[x(u, v), y(u, v), z(u, v)] \sqrt{EG - F^2} du dv$$

where R is the region corresponding to S in the uv -plane.

Similarly, if S is represented in the form $z = g(x, y)$, then from (10') and (10'') in the last section it follows that

$$(4) \quad \begin{aligned} \iint_S f(x, y, z) dA &= \iint_{\bar{S}} f[x, y, g(x, y)] \sqrt{1 + \left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2} dx dy \\ &= \iint_{\bar{S}} f[x, y, g(x, y)] \sec \gamma dx dy \quad (\gamma < \pi/2), \end{aligned}$$

where γ is the acute angle between the normal to S and the z -axis, and \bar{S} is the orthogonal projection of S in the xy -plane. Of course, here we must

assume that $g(x, y)$ is single-valued. Geometrically, this means that different points of S are projected onto different points of \bar{S} . The condition $\gamma < \pi/2$ means that the normal of S is nowhere perpendicular to the z -axis.

Example 1 (Moment of inertia). Find the moment of inertia I of a homogeneous spherical lamina $S: x^2 + y^2 + z^2 = a^2$ of mass M about the z -axis.

If a mass distributed over a surface S and $\mu(x, y, z)$ is the density of the mass (= mass per unit area), then the moment of inertia I of the mass with respect to a given axis L is defined by the surface integral

$$(5) \quad I = \iint_S \mu D^2 dA$$

where $D(x, y, z)$ is the distance of the point (x, y, z) from L .

Since, in the present example, μ is constant and S has the area $A = 4\pi a^2$ we have

$$\mu = \frac{M}{A} = \frac{M}{4\pi a^2}.$$

Representing S by (6), Sec. 6.5, we obtain from (4) in the previous section

$$\checkmark E = a^2 \cos^2 v, \quad F = 0, \quad G = a^2,$$

$$dA = \sqrt{EG - F^2} du dv = a^2 \cos v du dv.$$

Furthermore, the square of the distance of a point (x, y, z) , from the z -axis becomes

$$D^2 = x^2 + y^2 = a^2 \cos^2 v.$$

Hence we obtain the result

$$I = \iint_S \mu D^2 dA = \frac{M}{4\pi a^2} \int_{-\pi/2}^{\pi/2} \int_0^{2\pi} a^4 \cos^3 v du dv = \frac{Ma^2}{2} \int_{-\pi/2}^{\pi/2} \cos^3 v dv = \frac{2Ma^2}{3}.$$

In various applications there occur surface integrals for which the concept of orientation of a surface is essential. Therefore we shall now consider this concept, starting with the case of a smooth surface (cf. Sec. 6.5).

Let S be a smooth surface, and let P be any point of S . Then we may choose a unit normal vector n of S at P . The direction of n is then called the *positive normal direction* of S at P . Obviously there are two possibilities in choosing n .

A smooth surface S is said to be orientable if the positive normal direction, when given at an arbitrary point P_0 of S , can be continued in a unique and continuous way to the entire surface.

Hence the surface S is orientable provided there does not exist a closed curve C on S through P_0 such that the positive normal direction reverses when it is displaced continuously from P_0 along C and back to P_0 .

A sufficiently small portion of a smooth surface is always orientable. However, this may not hold in the large. There are nonorientable surfaces. A well-known example of such a surface is the Möbius⁶ strip shown in Fig. 192. When a normal vector, which is given at P_0 , is displaced continuously

⁶ AUGUST FERDINAND MÖBIUS (1790–1868), German mathematician, who made important contributions in the theory of surfaces and projective geometry.

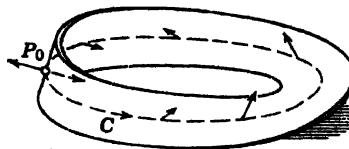
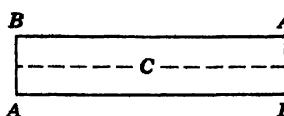


Fig. 192. Möbius strip.

along the curve C in Fig. 192, the resulting normal vector upon returning to P_0 is opposite to the original vector at P_0 . A model of a Möbius strip can be made by taking a long rectangular piece of paper and sticking the shorter sides together so that the two points A and the two points B in Fig. 192 coincide.

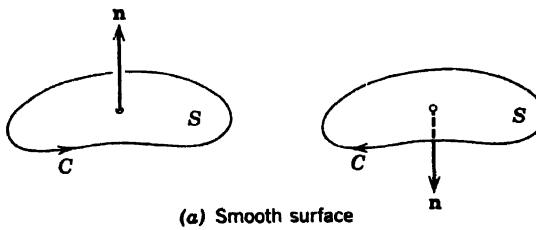
If a smooth surface S is orientable, then we may *orient* S by choosing one of the two possible directions of the normal vector \mathbf{n} .

If the boundary of S is a simple closed curve C , then we may associate with each of the two possible orientations of S an orientation of C , as shown in Fig. 193a. Using this simple idea, we may now readily extend the concept of orientation to piecewise smooth surfaces as follows.

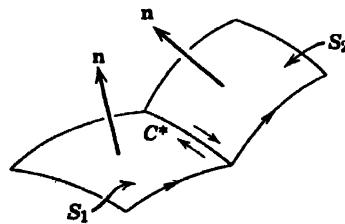
A piecewise smooth surface S is said to be *orientable* if we can orient each smooth piece of S in such a manner that along each curve C^* which is a common boundary of two pieces S_1 and S_2 the positive direction of C^* relative to S_1 is opposite to the positive direction of C^* relative to S_2 .

Figure 193b illustrates the situation for a surface consisting of two smooth pieces.

Let S be a given *orientable surface*. We orient S by choosing a unit normal vector \mathbf{n} . Denoting the angles between \mathbf{n} and the positive x , y , and z axes



(a) Smooth surface



(b) Piecewise smooth surface

Fig. 193. Orientation of a surface. *

by α , β , and γ , respectively, we have

$$(6) \quad \mathbf{n} = \cos \alpha \mathbf{i} + \cos \beta \mathbf{j} + \cos \gamma \mathbf{k}.$$

Let $u_1(x, y, z)$, $u_2(x, y, z)$, and $u_3(x, y, z)$ be given functions which are defined and continuous at every point of S . Then the integrals to be considered are usually written in the form

$$\iint_S u_1 dy dz, \quad \iint_S u_2 dz dx, \quad \iint_S u_3 dx dy,$$

and by definition this means

$$(7) \quad \begin{aligned} \iint_S u_1 dy dz &= \iint_S u_1 \cos \alpha dA, \\ \iint_S u_2 dz dx &= \iint_S u_2 \cos \beta dA, \\ \iint_S u_3 dx dy &= \iint_S u_3 \cos \gamma dA. \end{aligned}$$

It is clear that the value of such an integral depends on the choice of \mathbf{n} , that is, on the orientation of S . The transition to the opposite orientation corresponds to the multiplication of the integral by -1 , because then the components $\cos \alpha$, $\cos \beta$, and $\cos \gamma$ of \mathbf{n} are multiplied by -1 .

Sums of three such integrals may be written in a simple form by using vector notation. In fact, if we introduce the vector

$$\mathbf{u} = u_1 \mathbf{i} + u_2 \mathbf{j} + u_3 \mathbf{k},$$

then we obtain from the definition (7) the formula

$$(8) \quad \begin{aligned} &\iint_S (u_1 dy dz + u_2 dz dx + u_3 dx dy) \\ &= \iint_S (u_1 \cos \alpha + u_2 \cos \beta + u_3 \cos \gamma) dA = \iint_S \mathbf{u} \cdot \mathbf{n} dA. \end{aligned}$$

To evaluate the integrals (7) we may reduce them to double integrals over a plane region, as follows.

If S can be represented in the form $z = h(x, y)$ and is oriented such that \mathbf{n} points upward, then γ is acute. Hence from (10") in the last section and the above definition (7) we immediately have

$$(9a) \quad \iint_S u_3(x, y, z) dx dy = + \iint_R u_3[x, y, h(x, y)] dx dy$$

where R is the orthogonal projection of S in the xy -plane. If \mathbf{n} points downward, then γ is obtuse and we obtain

$$(9b) \quad \iint_S u_3(x, y, z) dx dy = - \iint_R u_3[x, y, h(x, y)] dx dy.$$

For the other two integrals in (7) the situation is quite similar.

If S is represented in parametric form

$$\mathbf{r}(u, v) = x(u, v)\mathbf{i} + y(u, v)\mathbf{j} + z(u, v)\mathbf{k},$$

then the normal vector is [cf. (3) in the last section]

$$(10) \quad (a) \quad \mathbf{n} = + \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|} \quad \text{or} \quad (b) \quad \mathbf{n} = - \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|}$$

depending on the orientation. Now from (10) and (11) in the last section it follows that

$$dA = \sqrt{EG - F^2} du dv = |\mathbf{r}_u \times \mathbf{r}_v| du dv.$$

Since $\cos \gamma = \mathbf{k} \cdot \mathbf{n}$, we thus obtain

$$\begin{aligned} \cos \gamma dA &= \mathbf{k} \cdot \mathbf{n} dA = \pm \mathbf{k} \cdot (\mathbf{r}_u \times \mathbf{r}_v) du dv = \pm \frac{x_u \quad y_u \quad z_u}{x_v \quad y_v \quad z_v} du dv \\ &= \pm \frac{\partial(x, y)}{\partial(u, v)} du dv, \end{aligned}$$

where the last expression involves the Jacobian (cf. Sec. 6.3). Hence in (7),

$$(11) \quad \iint_S u_3(x, y, z) dx dy = \pm \iint_R u_3[x(u, v), y(u, v), z(u, v)] \frac{\partial(x, y)}{\partial(u, v)} du dv$$

with the plus sign when S is oriented such that (10a) holds and the minus sign in the case of the opposite orientation. Here R is the region corresponding to S in the $u v$ -plane.

Problems

Representing S in parametric form and using (3), evaluate $\iint_S f(x, y, z) dA$ where

1. $f = x + 1$, $S: x^2 + y^2 = 1$, $0 \leq z \leq 3$
2. $f = 8x$, $S: z = x^2$, $0 \leq x \leq 2$, $-1 \leq y \leq 2$
3. $f = xy$, $S: z = xy$, $0 \leq x \leq 1$, $0 \leq y \leq 1$
4. $f = \arctan(y/x)$, $S: z = x^2 + y^2$, $1 \leq z \leq 4$
5. $f = 3x^3 \sin y$, $S: z = x^3$, $0 \leq x \leq 1$, $0 \leq y \leq \pi$
6. $f = x + y + z$, $S: x^2 + y^2 = 1$, $0 \leq z \leq 2$
7. $f = xy$, $S: x^2 + y^2 = 4$, $-1 \leq z \leq 1$

8. $f = \cos x + \sin y$, S : the portion of $x + y + z = 1$ in the first octant

9. $f = x(z^2 + 12y - y^4)$, S : $z = y^2$, $0 \leq x \leq 1$, $0 \leq y \leq 1$

10. $f = (x^2 + y^2)^2$, S : $z = (x^2 + y^2)^2$, $x^2 + y^2 \leq 1$

11. Justify the following formulas for the mass M and the center of gravity $(\bar{x}, \bar{y}, \bar{z})$ of a lamina S of density (mass per unit area) $\sigma(x, y, z)$ in space:

$$M = \iint_S \sigma \, dA, \quad \bar{x} = \frac{1}{M} \iint_S x\sigma \, dA, \quad \bar{y} = \frac{1}{M} \iint_S y\sigma \, dA, \quad \bar{z} = \frac{1}{M} \iint_S z\sigma \, dA.$$

12. Justify the following formulas for the moments of inertia of the lamina in Prob. 11 about the x -, y -, and z -axes, respectively:

$$I_x = \iint_S (y^2 + z^2)\sigma \, dA, \quad I_y = \iint_S (x^2 + z^2)\sigma \, dA, \quad I_z = \iint_S (x^2 + y^2)\sigma \, dA.$$

Find the moment of inertia of a lamina S of density 1 about an axis A where

13. S : $x^2 + y^2 = 1$, $0 \leq z \leq h$, A : the z -axis

14. S : $z^2 = x^2 + y^2$, $0 \leq z \leq h$, A : the z -axis

15. S : the torus in Ex. 3, Sec. 6.6, A : the z -axis

16. S as in Prob. 15, A : the line $x = a$ in the xz -plane

17. S as in Prob. 15, A : the line $x = a + b$ in the xz -plane

18. (Steiner's⁷ theorem) If I_A is the moment of inertia of a mass distribution of total mass M with respect to an axis A through the center of gravity, show that its moment of inertia I_B with respect to an axis B , which is parallel to A and has the distance k from it, is

$$I_B = I_A + k^2 M.$$

19. Using Steiner's theorem and the solution of Prob. 15, solve Probs. 16 and 17.

20. Construct a paper model of a Möbius strip. What happens if you cut it along the curve C in Fig. 192?

6.8 Triple Integrals. Divergence Theorem of Gauss

The triple integral is a generalization of the double integral introduced in Sec. 6.3. For defining this integral we consider a function $f(x, y, z)$ defined in a bounded closed⁸ region T of space. We subdivide T by planes parallel to the three coordinate planes. Then we number the parallelepipeds inside T from 1 to n . In each such parallelepiped we choose an arbitrary point, say, (x_k, y_k, z_k) in the k th parallelepiped, and form the sum

$$J_n = \sum_{k=1}^n f(x_k, y_k, z_k) \Delta V_k$$

⁷ JACOB STEINER (1796–1863), Swiss geometer.

⁸ Explained in footnote 2, p. 307 (with "sphere" instead of "circle").

where ΔV_k is the volume of the k th parallelepiped. This we do for larger and larger positive integers n in an arbitrary manner, but so that the lengths of the edges of the largest parallelepiped of subdivision approach zero as n approaches infinity. In this way, we obtain a sequence of real numbers J_{n_1}, J_{n_2}, \dots . We assume that $f(x, y, z)$ is continuous in T , and T is bounded by finitely many *smooth surfaces*.⁹ Then it can be shown (cf. Ref. [A1]) that the sequence converges to a limit which is independent of the choice of subdivisions and corresponding points (x_k, y_k, z_k) . This limit is called the *triple integral of $f(x, y, z)$ over the region T* , and is denoted by

$$\iiint_T f(x, y, z) dx dy dz \quad \text{or} \quad \iiint_R f(x, y, z) dV.$$

We shall now show that the triple integral of the divergence of a continuously differentiable vector function \mathbf{u} over a region T in space can be transformed into a surface integral of the normal component of \mathbf{u} over the boundary surface S of T . This can be done by means of the *divergence theorem which is the three-dimensional analogue of Green's theorem in the plane* considered in Sec. 6.4. The divergence theorem is of basic importance in various theoretical and practical considerations.

Divergence theorem of Gauss. *Let T be a closed bounded region in space whose boundary is a piecewise smooth¹⁰ orientable surface S . Let $\mathbf{u}(x, y, z)$ be a vector function which is continuous and has continuous first partial derivatives in some domain containing T . Then*

$$(1) \quad \iiint_T \operatorname{div} \mathbf{u} dV = \iint_S u_n dA = \iint_S (\mathbf{u} \cdot \mathbf{n}) dA$$

where

$$(2) \quad u_n = \mathbf{u} \cdot \mathbf{n}$$

is the component of \mathbf{u} in the direction of the outer normal of S with respect to T , and \mathbf{n} is the outer unit normal vector of S .

Remark. If we write \mathbf{u} and \mathbf{n} in terms of components, say

$$\mathbf{u} = u_1 \mathbf{i} + u_2 \mathbf{j} + u_3 \mathbf{k} \quad \text{and} \quad \mathbf{n} = \cos \alpha \mathbf{i} + \cos \beta \mathbf{j} + \cos \gamma \mathbf{k},$$

so that α , β , and γ are the angles between \mathbf{n} and the positive x , y , and z axes, respectively, the formula (1) takes the form

$$(3^*) \quad \begin{aligned} & \iiint_T \left(\frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z} \right) dx dy dz \\ &= \iint_S (u_1 \cos \alpha + u_2 \cos \beta + u_3 \cos \gamma) dA. \end{aligned}$$

⁹ Cf. Sec. 6.5.

¹⁰ Cf. Sec. 6.5.

Because of (8) in the last section this may be written

$$(3) \quad \iiint_T \left(\frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z} \right) dx dy dz = \iint_S (u_1 dy dz + u_2 dz dx + u_3 dx dy).$$

Proof of the divergence theorem. Clearly, (3*) is true, if the following three relations hold simultaneously:

$$(4) \quad \iiint_T \frac{\partial u_1}{\partial x} dx dy dz = \iint_S u_1 \cos \alpha dA,$$

$$(5) \quad \iiint_T \frac{\partial u_2}{\partial y} dx dy dz = \iint_S u_2 \cos \beta dA,$$

$$(6) \quad \iiint_T \frac{\partial u_3}{\partial z} dx dy dz = \iint_S u_3 \cos \gamma dA.$$

We first prove (6) for a *special region* T which is bounded by a piecewise smooth orientable surface S and has the property that any straight line parallel to any one of the coordinate axes and intersecting T has only *one* segment (or a single point) in common with T . This implies that T can be represented in the form

$$(7) \quad g(x, y) \leq z \leq h(x, y)$$

where (x, y) varies in the orthogonal projection \bar{R} of T in the xy -plane. Clearly, $z = g(x, y)$ represents the "bottom" S_2 of S (Fig. 194) whereas $z = h(x, y)$ represents the "top" S_1 of S , and there may be a remaining vertical portion S_3 of S . (The portion S_3 may degenerate into a curve, as for a sphere.)

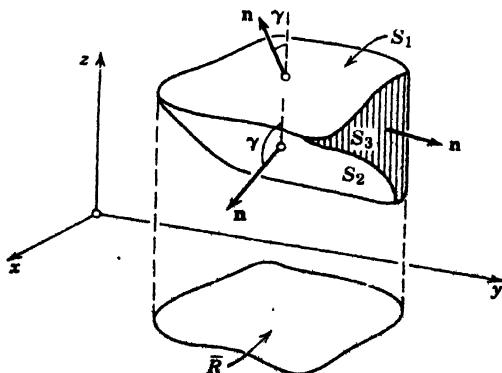


Fig. 194. Example of a special region.

To prove (6), we use (7). Since u is continuously differentiable in some domain containing T , we have

$$\iiint_T \frac{\partial u_3}{\partial z} dx dy dz = \iint_R \left[\int_g^{h(x,y)} \frac{\partial u_3}{\partial z} dz \right] dx dy.$$

We integrate the inner integral:

$$\int_g^h \frac{\partial u_3}{\partial z} dz = u_3(x, y, h) - u_3(x, y, g).$$

Hence

$$(8) \quad \begin{aligned} \iiint_T \frac{\partial u_3}{\partial z} dx dy dz \\ = \iint_R u_3[x, y, h(x, y)] dx dy - \iint_R u_3[x, y, g(x, y)] dx dy. \end{aligned}$$

We now show that the right-hand sides of (6) and (8) are equal. On the lateral portion S_3 of S (Fig. 194) we have $\gamma = \pi/2$ and $\cos \gamma = 0$. Hence this portion does not contribute to the surface integral in (6), and

$$\iint_S u_3 \cos \gamma dA = \iint_{S_1} u_3 \cos \gamma dA + \iint_{S_2} u_3 \cos \gamma dA$$

(cf. Fig. 194). On S_1 in Fig. 194 the angle γ is acute. Hence by (10") at the end of Sec. 6.6 we have $dA = \sec \gamma dx dy$. Since $\cos \gamma \sec \gamma = 1$, we thus obtain

$$\iint_{S_1} u_3 \cos \gamma dA = \iint_R u_3[x, y, h(x, y)] dx dy,$$

which is identical with the first double integral on the right-hand side of (8). On S_2 , the angle γ is obtuse (Fig. 194) so that $\pi - \gamma$ corresponds to the acute angle in (10"), Sec. 6.6, and, therefore,

$$dA = \sec(\pi - \gamma) dx dy = -\sec \gamma dx dy.$$

Hence

$$\iint_{S_2} u_3 \cos \gamma dA = - \iint_R u_3[x, y, g(x, y)] dx dy,$$

and this integral is identical with the last integral in (8). This proves (6).

The relations (4) and (5) then follow by merely relabelling the variables and using the fact that, by assumption, T has representations similar to (7), namely,

$$g(y, z) \leq x \leq h(y, z) \quad \text{and} \quad g^*(z, x) \leq y \leq h^*(z, x).$$

This establishes the divergence theorem for special regions.

For any region T which can be subdivided into *finitely many* special regions by means of auxiliary surfaces, the theorem follows by adding the result for each part separately; this procedure is analogous to that in the proof of Green's theorem in Sec. 6.4. The surface integrals over the auxiliary surfaces cancel in pairs, and the sum of the remaining surface integrals is the surface integral over the whole boundary S of T ; the volume integrals over the parts of T add up to that over T .

The divergence theorem is now proved for any bounded region which is of interest in practical problems. The extension to the most general region T of the type characterized in the theorem would require a certain limit process; this is similar to the situation in the case of Green's theorem in Sec. 6.4.

Green's theorem in the plane (Sec. 6.4) can be used for evaluating line integrals. Similarly, the divergence theorem is suitable for evaluating surface integrals, as may be illustrated by the following example.

Example 1. By transforming to a triple integral evaluate

$$I = \iint_S (x^3 dy dz + x^2 y dz dx + x^2 z dx dy)$$

where S is the closed surface consisting of the cylinder $x^2 + y^2 = a^2$ ($0 \leq z \leq b$) and the circular disks $z = 0$ and $z = b$ ($x^2 + y^2 \leq a^2$).

In (3) we now have

$$u_1 = x^3, \quad u_2 = x^2 y, \quad u_3 = x^2 z.$$

Hence, if we make use of the symmetry of the region T bounded by S , the triple integral in (3) assumes the form

$$\iiint_T (3x^2 + x^2 + x^2) dx dy dz = 4 \cdot 5 \int_0^b \int_0^a \int_0^{\sqrt{a^2 - y^2}} x^2 dx dy dz.$$

The integral over x equals $\frac{1}{3}(a^2 - y^2)^{3/2}$. Setting $y = a \cos t$, we have

$$dy = -a \sin t dt, \quad (a^2 - y^2)^{3/2} = a^3 \sin^3 t,$$

and the integral over y becomes

$$\frac{1}{3} \int_0^a (a^2 - y^2)^{3/2} dy = -\frac{1}{3} a^4 \int_{\pi/2}^0 \sin^4 t dt = \frac{\pi}{16} a^4.$$

The integral over z contributes the factor b , and therefore

$$I = 4 \cdot 5 \frac{\pi}{16} b a^4 = \frac{5\pi}{4} a^4 b.$$

6.9 Consequences and Applications of the Divergence Theorem

The divergence theorem has various applications and important consequences, some of which may be illustrated by the subsequent examples. In these examples, the regions and functions are assumed to satisfy the

conditions under which the divergence theorem holds, and in each case, \mathbf{n} is the outward unit normal vector of the boundary surface of the region, as before.

Example 1 (Representation of the divergence independent of the coordinates). Dividing both sides of (1) in Sec. 6.8 by the volume $V(T)$ of the region T , we obtain

$$(1) \quad \frac{1}{V(T)} \iiint_T \operatorname{div} \mathbf{u} dV = \frac{1}{V(T)} \iint_{S(T)} u_n dA$$

where $S(T)$ is the boundary surface of T . The basic properties of the triple integral are essentially the same as those of the double integral considered in Sec. 6.3. In particular the mean value theorem for triple integrals asserts that for any continuous function $f(x, y, z)$ in the region T under consideration,

$$\iiint_T f(x, y, z) dV = f(x_0, y_0, z_0)V(T)$$

where $Q: (x_0, y_0, z_0)$ is a (suitable) point in T . From this it follows that

$$(2) \quad \frac{1}{V(T)} \iiint_T \operatorname{div} \mathbf{u} dV = \operatorname{div} \mathbf{u}(x_0, y_0, z_0).$$

Let $P: (x_1, y_1, z_1)$ be any fixed point in T , and let T shrink down onto P , so that the maximum distance $d(T)$ of the points of T from P approaches zero. Then Q must approach P , and from (1) and (2) it follows that the divergence of \mathbf{u} at P is

$$(3) \quad \operatorname{div} \mathbf{u}(x_1, y_1, z_1) = \lim_{d(T) \rightarrow 0} \frac{1}{V(T)} \iint_{S(T)} u_n dA.$$

This formula is sometimes used as a definition of the divergence. While the definition of the divergence in Sec. 5.18 involves coordinates, the formula (3) is independent of the coordinate system. Hence from (3) it follows immediately that the divergence is a scalar function.

Example 2 (Physical interpretation of the divergence). By means of the divergence theorem we may obtain an intuitive interpretation of the divergence of a vector. For this purpose it is convenient to consider the flow of an incompressible fluid (cf. p. 295) of constant density $\rho = 1$ which is stationary or steady, that is, does not vary with time. Such a flow is determined by the field of its velocity vector $\mathbf{v}(P)$ at any point P .

Let S be the boundary surface of a region T in space, and let \mathbf{n} be the outward unit normal vector of S . The mass of fluid which flows through a small portion ΔS of S of area ΔA per unit time from the interior of S to the exterior is equal to $v_n \Delta A$ where¹¹ $v_n = \mathbf{v} \cdot \mathbf{n}$ is the normal component of \mathbf{v} in the direction of \mathbf{n} , taken at a suitable point of ΔS . Consequently, the total mass of fluid which flows across S from T to the outside per unit of time is given by the surface integral

$$\iint_S v_n dA.$$

¹¹ Note that v_n may be negative at a certain point, which means that fluid enters the interior of S at such a point.

Hence this integral represents the total flow out of T , and the integral

$$(4) \quad \frac{1}{V} \iint_S v_n dA,$$

where V is the volume of T , represents the average flow out of T . Since the flow is steady and the fluid is incompressible, the amount of fluid flowing outward must be continuously supplied. Hence, if the value of the integral (4) is different from zero, there must be *sources* (*positive sources and negative sources, called sinks*) in T , that is, points where fluid is produced or disappears.

If we let T shrink down to a fixed point P in T , we obtain from (4) the *source intensity* at P represented by the right-hand side of (3) (with u_n replaced by v_n). From this and (3) it follows that *the divergence of the velocity vector v of a steady incompressible flow is the source intensity of the flow at the corresponding point*. There are no sources in T if, and only if, $\operatorname{div} v \equiv 0$; in this case,

$$\iint_{S^*} v_n dA = 0$$

for any closed surface S^* in T .

/ Example 3 (Heat flow). We know that in a body heat will flow in the direction of decreasing temperature. Physical experiments show that the rate of flow is proportional to the gradient of the temperature. This means that the velocity v of the heat flow in a body is of the form

$$(5) \quad v = -K \operatorname{grad} U,$$

where $U(x, y, z, t)$ is the temperature, t is the time, and K is called the *thermal conductivity* of the body; in ordinary physical circumstances K is a constant.

Let R be a region in the body and let S be its boundary surface. Then the amount of heat leaving R per unit of time is

$$\iint_S v_n dA,$$

where $v_n = v \cdot n$ is the component of v in the direction of the outer unit normal vector n of S . This expression is obtained in a fashion similar to that in the preceding example. From (5) and the divergence theorem we obtain

$$(6) \quad \iint_S v_n dA = -K \iiint_R \operatorname{div} (\operatorname{grad} U) dx dy dz = -K \iiint_R \nabla^2 U dx dy dz;$$

cf. (5) in Sec. 5.18.

On the other hand, the total amount of heat H in R is

$$H = \iiint_R \sigma \rho U dx dy dz,$$

where the constant σ is the specific heat of the material of the body and ρ is its density (= mass per unit volume). Hence the time rate of decrease of H is

$$-\frac{\partial H}{\partial t} = -\iiint_R \sigma \rho \frac{\partial U}{\partial t} dx dy dz.$$

and this must be equal to the above amount of heat leaving R ; from (6) we thus have

$$-\iiint_R \sigma \rho \frac{\partial U}{\partial t} dx dy dz = -K \iiint_R \nabla^2 U dx dy dz$$

or

$$\iiint_R \left(\sigma \rho \frac{\partial U}{\partial t} - K \nabla^2 U \right) dx dy dz = 0.$$

Since this holds for any region R in the body, the integrand (if continuous) must be zero everywhere; that is,

$$(7) \quad \frac{\partial U}{\partial t} = c^2 \nabla^2 U \quad c^2 = \frac{K}{\sigma \rho}.$$

This partial differential equation is called the **heat equation**; it is fundamental for heat conduction. Methods for solving problems in heat conduction will be considered in Chap. 9.

Example 4 (A basic property of solutions of Laplace's equation). Consider the formula in the divergence theorem:

$$(8^*) \quad \iiint_T \operatorname{div} \mathbf{u} dV = \iint_S u_n dA.$$

Assume that \mathbf{u} is the gradient of a scalar function, say, $\mathbf{u} = \operatorname{grad} f$. Then [cf. (5) in Sec. 5.18]

$$\operatorname{div} \mathbf{u} = \operatorname{div} (\operatorname{grad} f) = \nabla^2 f.$$

Furthermore,

$$u_n = \mathbf{u} \cdot \mathbf{n} = \mathbf{n} \cdot \operatorname{grad} f,$$

and from (5) in Sec. 5.16 we see that the right side is the directional derivative of f in the outward normal direction of S . If we denote this derivative by $\partial f / \partial n$, our formula (8*) becomes

$$(8) \quad \iiint_T \nabla^2 f dV = \iint_S \frac{\partial f}{\partial n} dA.$$

Obviously this is the three-dimensional analogue of the formula (9) in Sec. 6.4.

Taking into account the assumptions under which the divergence theorem holds, we immediately obtain from (8) the following result.

Theorem 1. *Let $f(x, y, z)$ be a solution of Laplace's equation*

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0$$

in some domain D , and suppose that the second partial derivatives of f are continuous in D . Then the integral of the normal derivative of f over any piecewise smooth¹² closed orientable surface S in D is zero.

Example 5 (Green's theorem). Let f and g be scalar functions such that $\mathbf{u} = f \operatorname{grad} g$ satisfies the assumptions of the divergence theorem in some region T . Then

$$\operatorname{div} \mathbf{u} = \operatorname{div} (f \operatorname{grad} g) = f \nabla^2 g + \operatorname{grad} f \cdot \operatorname{grad} g$$

(cf. Prob. 12 at the end of Sec. 5.18). Furthermore

$$\mathbf{u} \cdot \mathbf{n} = \mathbf{n} \cdot (f \operatorname{grad} g) = f(\mathbf{n} \cdot \operatorname{grad} g).$$

¹² Cf. Sec. 6.5.

The expression $\mathbf{n} \cdot \operatorname{grad} g$ is the directional derivative of g in the direction of the outward normal vector \mathbf{n} of the surface S in the divergence theorem. If we denote this derivative by $\partial g / \partial n$, the formula in the divergence theorem becomes

$$(9) \quad \iiint_T (f \nabla^2 g + \operatorname{grad} f \cdot \operatorname{grad} g) dV = \iint_S f \frac{\partial g}{\partial n} dA.$$

This formula is called the **first Green's formula** or (together with the assumptions) the *first form of Green's theorem*.

By interchanging f and g we obtain a similar formula. On subtracting this formula from (9), we find

$$(10) \quad \iiint_T (f \nabla^2 g - g \nabla^2 f) dV = \iint_S \left(f \frac{\partial g}{\partial n} - g \frac{\partial f}{\partial n} \right) dA.$$

This formula is called the **second Green's formula** or (together with the assumptions) the *second form of Green's theorem*.

Example 6 (Uniqueness of solutions of Laplace's equation). Suppose that f satisfies the assumptions in Theorem 1 and is zero everywhere on a piecewise smooth closed orientable surface S in D . Then, setting $g = f$ in (9) and denoting the interior of S by T , we obtain

$$\iiint_T \operatorname{grad} f \cdot \operatorname{grad} f dV = \iiint_T |\operatorname{grad} f|^2 dV = 0.$$

Since by assumption $|\operatorname{grad} f|$ is continuous in T and on S and is non-negative, it must be zero everywhere in T . Hence $f_x = f_y = f_z = 0$, and f is constant in T and, because of continuity, equal to its value 0 on S . This proves

Theorem 2. *If a function $f(x, y, z)$ satisfies the assumptions of Theorem 1 and is zero at all points of a piecewise smooth closed orientable surface S in D , then it vanishes identically in the region T bounded by S .*

This theorem has an important consequence. Let f_1 and f_2 be functions which satisfy the assumptions of Theorem 1 and take on the same values on S . Then their difference $f_1 - f_2$ satisfies those assumptions and has the value 0 on S . Hence from Theorem 2 it follows that $f_1 - f_2 = 0$ throughout T , and we have the following result.

Theorem 3 (Uniqueness theorem). *Let f be a solution of Laplace's equation which has continuous second partial derivatives in a domain D , and let T be a region in D which satisfies the assumptions of the divergence theorem. Then f is uniquely determined in T by its values on the boundary surface S of T .*

Problems

Find the volume of the following regions by triple integration.

1. The tetrahedron with vertices $(0, 0, 0)$, $(2, 0, 0)$, $(0, 1, 0)$, $(0, 0, 3)$
2. The region in the first octant bounded by $y = x$, $y = x^2$, and $z = 1 - x$
3. The region between the paraboloid $z = 1 - x^2 - y^2$ and the xy -plane
4. The region bounded by the cylinders $x^2 + y^2 = 1$ and $y^2 + z^2 = 1$.

Find the total mass of a mass distribution of density σ in a region T where

5. $\sigma = xyz$, T : the cube $0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1$
6. $\sigma = x + y + 2z$, T : the tetrahedron with vertices $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$
7. $\sigma = xy$, T : the region in Prob. 6
8. $\sigma = z$, T : the region in the first octant bounded by $y = 1 - x^2$ and $z = x$

Find the moment of inertia

$$= \iiint_T (x^2 + y^2) dx dy dz$$

of a mass of density 1 in a region T about the z -axis, where T is

9. The cube $0 \leq x \leq c, 0 \leq y \leq c, 0 \leq z \leq c$
10. The cylinder $x^2 + y^2 \leq c^2, 0 \leq z \leq h$
11. The cone $x^2 + y^2 \leq z^2, 0 \leq z \leq h$
12. The interior of the sphere $x^2 + y^2 + z^2 = c^2$
13. Using the divergence theorem, show that the volume V of a region T bounded by a surface S is

$$V = \iint_S x dy dz = \iint_S y dz dx = \iint_S z dx dy = \frac{1}{3} \iint_S (x dy dz + y dz dx + z dx dy).$$

14. Verify the formulas in Prob. 13 for a cube.

15. Same task as in Prob. 14 for the cylinder $x^2 + y^2 \leq 1, 0 \leq z \leq h$.
16. Using $\mathbf{u} = xi + yj + zk$ in (1), Sec. 6.8, show that a region T with boundary surface S has the volume

$$\frac{1}{3} \iint_S r \cos \theta dA$$

where r is the distance of a variable point $P: (x, y, z)$ on S from the origin O and θ is the angle between the directed line OP and the outer normal to S at P .

17. Find the volume of a sphere of radius a by means of the formula in Prob. 16.

Evaluate the following surface integrals by the divergence theorem, assuming that S is oriented as in that theorem.

18. $\iint_S [(x+z) dy dz + (y+z) dz dx + (x+y) dx dy], S: x^2 + y^2 + z^2 = 4$
19. $\iint_S (x^2 dy dz + y^2 dz dx + z^2 dx dy), S$: the surface of the cube in Prob. 5
20. $\iint_S (e^x dy dz - ye^x dz dx + 3z dx dy), S$: the surface of the cylinder in Prob. 10
21. $\iint_S (yz dy dz + zx dz dx + xy dx dy), S$ as in Prob. 18
22. $\iint_S xyz dy dz, S$: the parallelepiped $0 \leq x \leq 3, 0 \leq y \leq 2, 0 \leq z \leq 1$

23. $\iint_S [\sin x \, dy \, dz + (2 - \cos x)y \, dz \, dx], \quad S \text{ as in Prob. 22}$

24. $\iint_S [(y \cos^2 x + y^3) \, dz \, dx + z(\sin^2 x - 3y^2) \, dx \, dy], \quad S \text{ as in Prob. 18}$

Let T be a closed bounded region in space and S its boundary surface. Using the divergence theorem, prove the following statements.

25. If g is harmonic¹³ in a domain containing T , then $\iint_S \frac{\partial g}{\partial n} \, dA = 0$.

26. If g is harmonic in a domain containing T , then $\iint_S g \frac{\partial g}{\partial n} \, dA = \iiint_T |\operatorname{grad} g|^2 \, dV$.

27. If g is harmonic in a domain containing T and $\partial g / \partial n = 0$ on S , then g is constant in T .

28. If f and g are harmonic in a domain containing T and $\partial f / \partial n = \partial g / \partial n$ on S , then $f = g + c$ in T , where c is a constant.

29. If f and g are harmonic in a domain containing T , then $\iint_S \left(f \frac{\partial g}{\partial n} - g \frac{\partial f}{\partial n} \right) \, dA = 0$.

30. Show that the Laplacian can be represented independently of all coordinate systems in the form

$$\nabla^2 f = \lim_{d(T) \rightarrow 0} \frac{1}{V(T)} \iint_{S(T)} \frac{\partial f}{\partial n} \, dA$$

where $d(T)$ is the maximum distance of the points of a region T bounded by $S(T)$ from the point at which the Laplacian is evaluated and $V(T)$ is the volume of T . Hint: put $u = \operatorname{grad} f$ in (3) and use (5), Sec. 5.16, with $b = n$, the outer unit normal vector to S .

6.10 Stokes's Theorem

In Sec. 6.4 it was shown that double integrals over a plane region can be transformed into line integrals over the boundary curve of the region. Generalizing this result, we shall now consider the corresponding problem in the case of a surface integral.

Stokes's theorem.¹⁴ Let S be a piecewise smooth¹⁵ oriented surface in space and let the boundary of S be a piecewise smooth simple closed curve C . Let $\mathbf{v}(x, y, z)$ be a continuous vector function which has continuous first partial

¹³ This means that g is a solution of Laplace's equation, and its second partial derivatives are continuous in T .

¹⁴ GEORGE GABRIEL STOKES (1819–1903). Irish mathematician and physicist, who made important contributions to the theory of infinite series and several branches of theoretical physics.

¹⁵ Cf. Secs. 6.1 and 6.5.

derivatives in a region of space which contains S in its interior. Then

$$(1) \quad \iint_S (\operatorname{curl} \mathbf{v})_n \, dA = \int_C v_t \, ds;$$

here $(\operatorname{curl} \mathbf{v})_n = (\operatorname{curl} \mathbf{v}) \cdot \mathbf{n}$ is the component of $\operatorname{curl} \mathbf{v}$ in the direction of a unit normal vector \mathbf{n} of S ; the integration around C is taken in the sense shown in Fig. 195, and v_t is the component of \mathbf{v} in the direction of the tangent vector of C in Fig. 195.

Proof. We first prove Stokes's theorem for a surface S which can be represented simultaneously in the forms

$$(2) \quad (a) z = f(x, y), \quad (b) y = g(x, z), \quad \text{and} \quad (c) x = h(y, z),$$

where f , g , and h are continuous functions of the respective variables and have continuous first partial derivatives. Let

$$(3*) \quad \mathbf{n} = \cos \alpha \mathbf{i} + \cos \beta \mathbf{j} + \cos \gamma \mathbf{k}$$

be the "upper" unit normal vector of S (Fig. 196), and let

$$\mathbf{v} = v_1 \mathbf{i} + v_2 \mathbf{j} + v_3 \mathbf{k}.$$

If we represent C in the form $\mathbf{r} = \mathbf{r}(s)$ where the arc length s increases in the direction of integration, the unit tangent vector is

$$\frac{d\mathbf{r}}{ds} = \frac{dx}{ds} \mathbf{i} + \frac{dy}{ds} \mathbf{j} + \frac{dz}{ds} \mathbf{k},$$

and, therefore,

$$v_t = \mathbf{v} \cdot \frac{d\mathbf{r}}{ds} = v_1 \frac{dx}{ds} + v_2 \frac{dy}{ds} + v_3 \frac{dz}{ds}$$

From this we obtain

$$\mathbf{v} \cdot \frac{d\mathbf{r}}{ds} \, ds = \mathbf{v} \cdot \frac{d\mathbf{r}}{ds} \, ds = v_1 \, dx + v_2 \, dy + v_3 \, dz.$$

Consequently, if we use the representation of the curl in terms of right-handed Cartesian coordinates [cf. (1), Sec. 5.19], the formula in Stokes's theorem

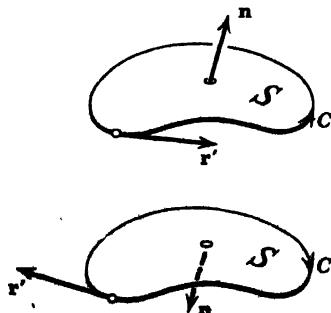


Fig. 195. Stokes's theorem.

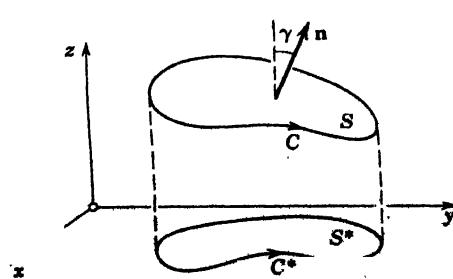


Fig. 196. Proof of Stokes's theorem.

may be written

$$(3) \quad \iint_S \left[\left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z} \right) \cos \alpha + \left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x} \right) \cos \beta + \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right) \cos \gamma \right] dA \\ = \int_C (v_1 dx + v_2 dy + v_3 dz),$$

where α , β , and γ are defined in connection with (3*).

We prove that in (3) the integrals over the terms involving v_1 are equal; that is,

$$(4) \quad \iint_S \left(\frac{\partial v_1}{\partial z} \cos \beta - \frac{\partial v_1}{\partial y} \cos \gamma \right) dA = \int_C v_1 dx.$$

Let S^* be the orthogonal projection of S in the xy -plane and let C^* be its boundary which is oriented as shown in Fig. 196. Using the representation (2a) of S , we may write the line integral over C as a line integral over C^* :

$$\int_C v_1(x, y, z) dx = \int_{C^*} v_1[x, y, f(x, y)] dx.$$

We now apply Green's theorem in the plane (Sec. 6.4) to the functions $v_1[x, y, f(x, y)]$ and 0 [instead of f and g in Sec. 6.4]. Then

$$\int_{C^*} v_1[x, y, f(x, y)] dx = - \iint_{S^*} \frac{\partial v_1}{\partial y} dx dy.$$

In the integral on the right,

$$\frac{\partial v_1[x, y, f(x, y)]}{\partial y} = \frac{\partial v_1(x, y, z)}{\partial y} + \frac{\partial v_1(x, y, z)}{\partial z} \frac{\partial f}{\partial y} \quad [z = f(x, y)]$$

and, therefore,

$$(5) \quad \int_C v_1(x, y, z) dx = - \iint_{S^*} \left(\frac{\partial v_1}{\partial y} + \frac{\partial v_1}{\partial z} \frac{\partial f}{\partial y} \right) dx dy.$$

We prove that the integral on the right-hand side of (5) equals the integral on the left-hand side of (4). In the latter integral, we introduce x and y as variables of integration. Writing (2a) in the form

$$F(x, y, z) = z - f(x, y) = 0,$$

we obtain

$$\text{grad } F = - \frac{\partial f}{\partial x} \mathbf{i} - \frac{\partial f}{\partial y} \mathbf{j} + \mathbf{k}.$$

Denoting the length of $\text{grad } F$ by a , we have

$$|\text{grad } F| = \sqrt{1 + \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2}.$$

Since $\text{grad } F$ is normal to S , we obtain

$$\mathbf{n} = \pm \frac{\text{grad } F}{a}$$

where \mathbf{n} is defined as before. But the components of both \mathbf{n} and $\text{grad } F$ in the positive z -direction are positive, and, therefore,

$$\mathbf{n} = + \frac{\text{grad } F}{a}.$$

Hence from the preceding representations of \mathbf{n} and $\text{grad } F$ in terms of components with respect to the xyz -coordinate system we see that

$$\cos \alpha = -\frac{1}{a} \frac{\partial f}{\partial x}, \quad \cos \beta = -\frac{1}{a} \frac{\partial f}{\partial y}, \quad \cos \gamma = \frac{1}{a}.$$

Furthermore, from (10'') at the end of Sec. 6.6 it follows that in (4), $dA = a dx dy$. Hence

$$\iint_S \left(\frac{\partial v_1}{\partial z} \cos \beta - \frac{\partial v_1}{\partial y} \cos \gamma \right) dA = \iint_S \left(\frac{\partial v_1}{\partial z} \left(-\frac{1}{a} \frac{\partial f}{\partial y} \right) - \frac{\partial v_1}{\partial y} \frac{1}{a} \right) a dx dy.$$

The integral on the right equals the right-hand side of (5), and (4) is proved.

If $-\mathbf{n}$ were chosen for the positive normal direction, then by assumption the sense of integration along C would be reversed and the result would be the same as before. This shows that (4) holds for both choices of the positive normal of S .

Using the representations (2b) and (2c) of S and reasoning exactly as before, we obtain

$$(6) \quad \iint_S \left(\frac{\partial v_2}{\partial x} \cos \gamma - \frac{\partial v_2}{\partial z} \cos \alpha \right) dA = \int_C v_2 dy$$

and

$$(7) \quad \iint_S \left(\frac{\partial v_3}{\partial y} \cos \alpha - \frac{\partial v_3}{\partial x} \cos \beta \right) dA = \int_C v_3 dz.$$

By adding (4), (6), and (7) we obtain (1). This proves Stokes's theorem for a surface S which can be represented simultaneously in the forms (2a), (2b), and (2c).

As in the proof of the divergence theorem, our result may be immediately extended to a surface S which can be decomposed into finitely many pieces, each of which is of the type considered before. This covers most of the cases which are of practical interest. The proof in the case of a most general surface S satisfying the assumptions of the theorem would require a limit process; this is similar to the situation in the case of Green's theorem in Sec. 6.4.

6.11 Consequences and Applications of Stokes's Theorem

Example 1 (Green's theorem in the plane as a special case of Stokes's theorem). Let $\mathbf{v} = v_1 \mathbf{i} + v_2 \mathbf{j}$ be a vector function which is continuously differentiable in a simply-connected bounded closed region S in the xy -plane whose boundary C is a piecewise smooth simple closed curve. Then, according to (1) in Sec. 5.19,

$$(\operatorname{curl} \mathbf{v})_n = \frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}.$$

Furthermore, $v_t ds = v_1 dx + v_2 dy$, and (1) in the last section takes the form

$$\iint_S \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right) dA = \int_C (v_1 dx + v_2 dy).$$

This shows that Green's theorem in the plane (Sec. 6.4) is a special case of Stokes's theorem.

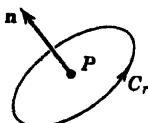


Fig. 197.
Example 2.

Example 2 (Physical interpretation of the curl). Let S_r be a circular disk of radius r and center P bounded by the circle C_r (Fig. 197), and let $\mathbf{v}(Q) \equiv \mathbf{v}(x, y, z)$ be a continuously differentiable vector function in S_r . Then by Stokes's theorem and the law of the mean for integrals,

$$\int_{C_r} v_t ds = \iint_{S_r} (\operatorname{curl} \mathbf{v})_n dA = [\operatorname{curl} \mathbf{v}(P^*)]_n A_r$$

where A_r is the area of S_r , and P^* is a suitable point of S_r . This may be written

$$[\operatorname{curl} \mathbf{v}(P^*)]_n = \frac{1}{A_r} \int_{C_r} v_t ds.$$

In the case of a fluid motion with velocity \mathbf{v} , the integral

$$\int_{C_r} v_t ds$$

is called the **circulation** of the flow around C_r ; it measures the extent to which the corresponding fluid motion is a rotation around the circle C_r . If we now let r approach zero, we find

$$(1) \quad [\operatorname{curl} \mathbf{v}(P)]_n = \lim_{r \rightarrow 0} \frac{1}{A_r} \int_{C_r} v_t ds;$$

that is, the component of the curl in the positive normal direction can be regarded as the **specific circulation** (circulation per unit area) of the flow in the surface at the corresponding point.

Problems

Evaluate $\iint_S (\operatorname{curl} \mathbf{v})_n dA$, where

1. $\mathbf{v} = xi + yj$; S : the square $0 \leq x \leq 1, 0 \leq y \leq 1, z = 1$

2. $\mathbf{v} = z\mathbf{i} + x\mathbf{j} + y\mathbf{k}$, S as in Prob. 1
3. $\mathbf{v} = -y^3\mathbf{i} + x^3\mathbf{j}$, S : the circular disk $x^2 + y^2 \leq 1$, $z = 0$
4. Verify Stokes's theorem for \mathbf{v} and S in Prob. 1.
5. Verify Stokes's theorem for \mathbf{v} and S in Prob. 3.
6. Show that if \mathbf{v} and S satisfy the assumptions of Stokes's theorem and $\mathbf{v} = \operatorname{grad} f$, then $\int_C v_t ds = 0$, where C is the boundary of S .

Evaluate $\int_C v_t ds$ by Stokes's theorem, where, with respect to right-handed Cartesian coordinates,

7. $\mathbf{v} = 2yi + zj + 3yk$, C : the intersection of $x^2 + y^2 + z^2 = 6z$ and $z = x + 3$, oriented clockwise as viewed from the origin
8. $\mathbf{v} = -3yi + 3xj + zk$, C : the circle $x = \cos \alpha$, $y = \sin \alpha$, $z = 1$ ($0 \leq \alpha \leq 2\pi$)
9. $\mathbf{v} = y^2\mathbf{i} + x^2\mathbf{j} - (x + z)\mathbf{k}$, C : the boundary of the triangle with vertices at $(0, 0, 0)$, $(1, 0, 0)$, $(1, 1, 0)$, (counterclockwise)
10. $\mathbf{v} = 4zi - 2xj + 2zk$, C : the intersection of $x^2 + y^2 = 1$ and $z = y + 1$, oriented clockwise as viewed from the origin
11. $\mathbf{v} = x^2\mathbf{i} + y^2\mathbf{j} + z^2\mathbf{k}$, C : the intersection of $x^2 + y^2 + z^2 = a^2$ and $z = y^2$

Using (3) in the last section, evaluate the following line integrals in the clockwise sense (as viewed from the origin) around the given contour C .

12. $\int_C (z dx + x dy + y dz)$, C : the boundary of the triangle with vertices $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$
13. $\int_C (yz dx + xz dy + xy dz)$, C : the intersection of $x^2 + y^2 = 1$ and $z = y^2$
14. $\int_C (\sin z dx - \cos x dy + \sin y dz)$, C : the boundary of the rectangle $0 \leq x \leq \pi$, $0 \leq y \leq 1$, $z = 3$
15. $\int_C (e^x dx + 2y dy - dz)$, C : $x^2 + y^2 = 4$, $z = 2$

6.12 Line Integrals Independent of Path

In Sec. 6.2 we saw that the value of a line integral

$$(1) \quad \int_C (f dx + g dy + h dz)$$

will in general depend not only on the end points P and Q of the path C but also on C ; that is, if we integrate from P to Q along different paths, we shall, in general, obtain different values of the integral. We shall now see under what conditions that value depends only on P and Q but does not depend on the path C from P to Q . This problem is of great importance. We first state the following definition.

Let $f(x, y, z)$, $g(x, y, z)$, and $h(x, y, z)$ be functions which are defined and continuous in a domain D of space. Then a line integral of the form (1) is said to be **independent of path in D** , if for every pair of end points P and Q in D the value of the integral is the same for all paths C in D starting from P and ending at Q . This value will then, in general, depend on the choice of P and Q but not on the choice of the path joining them.

For formulating our results it will be convenient to use the following concepts.

An expression of the form

$$(2) \quad f dx + g dy + h dz,$$

where f , g , h are functions defined in a domain D in space, is called a *first-order differential form* in three variables. This form is said to be **exact** or an *exact differential in D* , if it is the differential

$$(3) \quad du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy + \frac{\partial u}{\partial z} dz$$

of a *single-valued* differentiable function $u(x, y, z)$ everywhere in D ; that is,

$$(4) \quad f dx + g dy + h dz = du.$$

By comparing (3) and (4) we see that *the form (2) is exact in D if, and only if, there is a single-valued differentiable function $u(x, y, z)$ such that*

$$(5) \quad f = \frac{\partial u}{\partial x}, \quad g = \frac{\partial u}{\partial y}, \quad h = \frac{\partial u}{\partial z}$$

everywhere in D .

In vector language this means that *the form (2) is exact in D if, and only if, the vector function*

$$\mathbf{v} = f\mathbf{i} + g\mathbf{j} + h\mathbf{k}$$

is the gradient of a single-valued function $u(x, y, z)$ in D :

$$(5') \quad \mathbf{v} = \text{grad } u = \frac{\partial u}{\partial x} \mathbf{i} + \frac{\partial u}{\partial y} \mathbf{j} + \frac{\partial u}{\partial z} \mathbf{k}.$$

Theorem 1. *Let $f(x, y, z)$, $g(x, y, z)$, and $h(x, y, z)$ be continuous in a domain D in space. Then the line integral*

$$(6) \quad \int (f dx + g dy + h dz)$$

is independent of path in D if and only if the differential form under the integral sign is exact in D .

Proof. (a) Suppose that the integral under consideration is independent of path in D . We choose any fixed point $P: (x_0, y_0, z_0)$ and a point $Q: (x, y, z)$

in D . Then we define a function $u(x, y, z)$ as follows:

$$(7) \quad u(x, y, z) = u_0 + \int_C (f dx + g dy + h dz),$$

where u_0 is a constant and we integrate along an arbitrary path in D from P to Q . Since the integral is independent of path and P is fixed, it does indeed depend only on the coordinates x, y, z of the end point Q and defines a function $u(x, y, z)$ in D . It remains to show that from (7) the relations (5) follow, which means that $f dx + g dy + h dz$ is exact in D . We prove the first of these three relations. Since the integral is independent of path, we may integrate from P to a point $Q_1(x_1, y_1, z_1)$ and then parallel to the x -axis along the line segment from Q_1 to Q (Fig. 198); here Q_1 is chosen so that the whole segment $Q_1 Q$ lies in D . Then



Fig. 198. Proof of Theorem 1.

$$u(x, y, z) = u_0 + \int_P^{Q_1} (f dx + g dy + h dz) + \int_{Q_1}^Q (f dx + g dy + h dz).$$

We take the partial derivative with respect to x . Since P and Q_1 do not depend on x , that derivative of the first integral is zero. Since on the segment $Q_1 Q$, y and z are constant, the last integral may be written as the definite integral

$$\int_{x_1}^x f(x, y, z) dx.$$

Hence the partial derivative of that integral with respect to x is $f(x, y, z)$, and the first of the relations (5) is proved. The other two relations can be proved by the same argument.

(b) Conversely, suppose that $f dx + g dy + h dz$ is exact in D . Then (5) holds in D for some single-valued function u . Let C be any path from P to Q in D , and let

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k} \quad t \leq t_1$$

be a parametric representation of C so that P corresponds to $t = t_0$ and Q corresponds to $t = t_1$. Then

$$\begin{aligned} \int_P^Q (f dx + g dy + h dz) &= \int_P^Q \left(\frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy + \frac{\partial u}{\partial z} dz \right) \\ &= \int_{t_0}^{t_1} \frac{du}{dt} dt = u[x(t), y(t), z(t)]_{t=t_0}^{t=t_1} = u(Q) - u(P). \end{aligned}$$

This shows that the value of the integral is simply the difference of the values of u at the two end points of C and is, therefore, independent of the path C . Theorem 1 is proved.

The last formula in the proof,

$$(8) \quad \int_P^Q (f dx + g dy + h dz) = u(Q) - u(P),$$

is the analogue of the usual formula

$$\int_a^b f(x) dx = F(x) \Big|_a^b = F(b) - F(a) \quad [\text{where } F'(x) = f(x)]$$

for evaluating definite integrals which is known from elementary calculus. Formula (8) should be used whenever a line integral is independent of path. The practical use of this formula will be explained at the end of this section.

We remember that the work W done by a (variable) force

$$\mathbf{p} = f\mathbf{i} + g\mathbf{j} + h\mathbf{k}$$

in the displacement of a particle along a path C in space is given by the line integral (7) in Sec. 6.2,

$$W = \int_C \mathbf{p} \cdot d\mathbf{r} = \int_C (f dx + g dy + h dz),$$

the integration being taken in the sense of the displacement. From Theorem 1 we conclude that W depends only on the end points P and Q of the path C if, and only if, the differential form under the integral sign is exact; and this form is exact if, and only if, \mathbf{p} is the gradient of a scalar function u . In this case the field of force given by \mathbf{p} is said to be **conservative** (cf. also at the end of Sec. 5.16).

Example 1. Find the work W done by the force $\mathbf{p} = yz\mathbf{i} + xz\mathbf{j} + xy\mathbf{k}$ in the displacement of a particle along the straight segment C from $P: (1, 1, 1)$ to $Q: (3, 3, 2)$. The work is given by the integral

$$W = \int_C (yz dx + xz dy + xy dz).$$

This integral is of the form (1) where

$$f = yz, \quad g = xz, \quad h = xy,$$

and (5) holds for $u = xyz$. Hence W is independent of path and we may use (8), finding

$$W = u(Q) - u(P) = u(3, 3, 2) - u(1, 1, 1) = 18 - 1 = 17.$$

From Theorem 1 we may immediately derive the following important result.

Theorem 2. Let f, g, h be continuous in a domain D of space. Then the line integral

$$\int_C (f dx + g dy + h dz)$$

is independent of path in D if, and only if, it is zero on every simple closed path in D .

:

Proof. (a) Let C be a simple closed path in D and suppose that the integral is independent of path in D . We subdivide C into two arcs C_1 and C_2 (Fig. 199). Then

$$(9) \quad \oint_C (f dx + g dy + h dz) = \int_{C_1} (f dx + g dy + h dz) + \int_{C_2} (f dx + g dy + h dz).$$

Because of independence of path,

$$\int_{C_1} (f dx + g dy + h dz) = \int_{C_1^*} (f dx + g dy + h dz) = - \int_{C_1} (f dx + g dy + h dz),$$

where C_1^* denotes C_1 traced in the reverse direction.

From this it follows that the integral on the left-hand side of (9) is zero.

(b) Conversely, suppose that the integral under consideration is zero on every simple closed path in D . Let P and Q be any two points of D , and let C_1 and C_2 be two paths in D which join P and Q and do not cross (Fig. 199). C_1 and C_2 together form a simple closed path C . Therefore,

$$\int_{C_1} (f dx + g dy + h dz) + \int_{C_2} (f dx + g dy + h dz) = \oint_C (f dx + g dy + h dz) = 0.$$

From this we obtain

$$\begin{aligned} \int_{C_1} (f dx + g dy + h dz) &= - \int_{C_1} (f dx + g dy + h dz) = \int_{C_1^*} (f dx + g dy + h dz). \end{aligned}$$

This completes the proof of Theorem 2.

The extension of this theorem to the case of closed paths crossing a finite number of times is immediate.

To make effective practical use of the simple Theorem 1 we need a criterion by which we can decide whether the differential form under the integral sign is exact. Such a criterion will be given in Theorem 3, below. For formulating this theorem we shall use the following concept.

A domain D is called **simply connected** if every closed curve in D can be continuously shrunk to any point in D without leaving D .

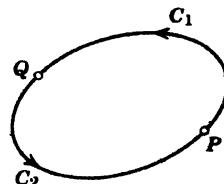


Fig. 199. Proof of Theorem 2.

For example, the interior of a sphere or a cube, the interior of a sphere with finitely many points removed, and the domain between two concentric spheres are simply connected, while the interior of a torus (cf. Sec. 6.6) and the interior of a cube with one space diagonal removed are not simply connected.

Theorem 3. Let $f(x, y, z)$, $g(x, y, z)$, and $h(x, y, z)$ be continuous functions having continuous first partial derivatives in a domain D in space. If the line integral

$$(10) \quad \int (f dx + g dy + h dz)$$

is independent of path in D (and, consequently, $f dx + g dy + h dz$ is exact in D), then

$$(11) \quad \frac{\partial h}{\partial y} = \frac{\partial g}{\partial z}, \quad \frac{\partial f}{\partial z} = \frac{\partial h}{\partial x}, \quad \frac{\partial g}{\partial x} = \frac{\partial f}{\partial y}$$

or, in vector language,

$$(11') \quad \operatorname{curl} \mathbf{v} = \mathbf{0} \quad (\mathbf{v} = f\mathbf{i} + g\mathbf{j} + h\mathbf{k})$$

everywhere in D . Conversely, if D is simply connected and (11) holds everywhere in D , then the integral (10) is independent of path in D (and, consequently, $f dx + g dy + h dz$ is exact in D).

Proof. (a) Suppose that (10) is independent of path in D . Then, by Theorem 1, the form $f dx + g dy + h dz$ is exact in D and, by (5'),

$$\mathbf{v} = f\mathbf{i} + g\mathbf{j} + h\mathbf{k} = \operatorname{grad} u.$$

From this and (3) in Sec. 5.19 we obtain

$$\operatorname{curl} \mathbf{v} = \operatorname{curl} (\operatorname{grad} u) = \mathbf{0}.$$

(b) Conversely, suppose that D is simply connected and (11') holds everywhere in D . Let C be any simple closed path in D . Since D is simply connected, we can find a surface S in D bounded by C . Stokes's theorem (Sec. 6.10) is applicable, and we obtain

$$\int_C (f dx + g dy + h dz) = \int_C v_t ds = \iint_S (\operatorname{curl} \mathbf{v})_n dA = 0$$

for proper direction on C and normal \mathbf{n} on S . From this and Theorem 2 it follows that the integral (10) is independent of path in D . This completes the proof.

We see that in the case of a line integral

$$\int (f dx + g dy)$$

in the xy -plane, (11) reduces to the one relation

$$(11^*) \quad \frac{\partial g}{\partial x} = \frac{\partial f}{\partial y}.$$

The assumption that D be simply connected is essential and cannot be omitted. This can be seen from the following example.

Example 2. Let

$$f = -\frac{y}{x^2 + y^2}, \quad g = \frac{x}{x^2 + y^2}, \quad h = 0.$$

Differentiation shows that (11*) is satisfied in any domain of the xy -plane not containing the origin, for example, in the domain $D: \frac{1}{2} < \sqrt{x^2 + y^2} < \frac{3}{2}$ shown in Fig. 200. Clearly, D is not simply connected. If the integral

$$I = \int (f dx + g dy) = \int \frac{-y dx + x dy}{x^2 + y^2}$$

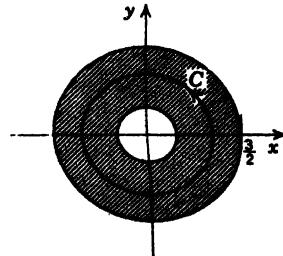


Fig. 200. Example 2.

were independent of path in D , then $I = 0$ on any closed curve in D , for example, on the circle $x^2 + y^2 = 1$. But setting $x = r \cos \theta$, $y = r \sin \theta$, and noting that the circle is represented by $r = 1$, we easily obtain

$$I = \int_0^{2\pi} d\theta = 2\pi,$$

the integration being taken once around the circle in the counterclockwise sense. In fact, since D is not simply connected, we cannot apply Theorem 3 and conclude that I is independent of path in D . Furthermore, we see that

$$fdx + g dy = du \quad \text{where} \quad u = \operatorname{arc tan} \frac{y}{x} = \theta,$$

but u is not single-valued in D . On the other hand, if we take the so-called "principal value" of u , defined by $-\pi < u \leq \pi$, then u is not differentiable (not even continuous) at the points of the negative x -axis in D ; this was required in connection with the exactness of a differential form.

If a line integral is independent of path, it can be evaluated by means of the formula (8), and the corresponding function $u(x, y, z)$ can be found by a single integration as follows.

Let $P: (x_0, y_0, z_0)$ and $Q: (x, y, z)$, as before. Then we obtain from (8)

$$u(Q) = u(x, y, z) = u(x_0, y_0, z_0) + \int_{(x_0, y_0, z_0)}^{(x, y, z)} (f dx + g dy + h dz);$$

integrating from P to Q along the line segment

$$(12) \quad x = x_0 + (x - x_0)t, \quad y = y_0 + (y - y_0)t, \quad z = z_0 + (z - z_0)t$$

(where $0 \leq t \leq 1$), we obtain $dx = (x - x_0) dt$, etc. Therefore,

$$(13) \quad u(x, y, z) = u(x_0, y_0, z_0) + \int_0^1 [(x - x_0)f + (y - y_0)g + (z - z_0)h] dt$$

where $f = f(x, y, z)$, etc., and x, y, z are given by (12).

Another way of determining $u(x, y, z)$ is to integrate one of the relations (5), for example, $\partial u / \partial x = f$, finding

$$u = \int f dx + a(y, z),$$

and then to determine the function $a(y, z)$ by differentiating this result with respect to y and z and using the other two relations (5); cf. the following example.

Example 3. Evaluate

$$I = \int_C [2xyz^2 dx + (x^2z^2 + z \cos yz) dy + (2x^2yz + y \cos yz) dz]$$

on any path C from $P: (0, 0, 1)$ to $Q: (1, \pi/4, 2)$.

From Theorem 3 it follows that the integral is independent of path in space. Since $f = 2xyz^2$, we obtain from the first relation (5)

$$(14) \quad u = \int f dx = x^2yz^2 + a(y, z).$$

From this and the second relation (5) it follows that

$$\frac{\partial u}{\partial y} = x^2z^2 + \frac{\partial a}{\partial y} = g = x^2z^2 + z \cos yz.$$

Hence we must have

$$\frac{\partial a}{\partial y} = z \cos yz, \quad a = \sin yz + c(z).$$

From this, the third relation (5), and (14) we obtain

$$\frac{\partial u}{\partial z} = 2x^2yz + y \cos yz + \frac{dc}{dz} = h = 2x^2yz + y \cos yz.$$

Hence $dc/dz = 0$, $c = \text{const}$ and, therefore,

$$u(x, y, z) = x^2yz^2 + a = x^2yz^2 + \sin yz + c.$$

This yields the result

$$I = [x^2yz^2 + \sin yz + c] \Big|_P^Q = \pi + \sin \frac{\pi}{2} = \pi + 1.$$

Problems

Are the following forms exact?

- | | |
|----------------------------------|--|
| 1. $x dx - y dy + z dz$ | 2. $x^2 dx + 2xy dy + y^2 dz$ |
| 3. $e^x dx + e^y dy + e^z dz$ | 4. $(y + z \cos x) dx + x dy + \sin x dz$ |
| 5. $z \cos yz dy + y \cos yz dz$ | 6. $y^2 dx + x^2 dy + x^3 dz$ |
| 7. $yz dx + zx dy + xy dz$ | 8. $y^2 z^2 dx + 2xyz^2 dy + 3xy^2 z^2 dz$ |

In each case show that the given differential form is exact and find a function u such that the form equals du .

9. $x \, dx - y \, dy - z \, dz$

10. $dx + dy + dz$

11. $dx + z \, dy + y \, dz$

12. $(z^2 - 2xy) \, dx - x^2 \, dy + 2xz \, dz$

13. $\cos x \, dx - 2yz \, dy - y^2 \, dz$

14. $6xy \, dx + 3x^2 \, dy - 4z \, dz$

Show that the form under the integral sign is exact and evaluate:

15. $\int_{(0,0,0)}^{(1,1,1)} (x \, dx + y \, dy - z \, dz)$

16. $\int_{(1,0,0)}^{(2,1,4)} [yz \, dx + (xz + 1) \, dy + xy \, dz]$

17. $\int_{(0,-1,-1)}^{(1,3,3)} (\cos x \, dx + z \, dy + y \, dz)$

18. $\int_{(1,3,2)}^{(4,0,1)} [z \, dx - z \, dy + (x - y) \, dz]$

19. $\int_{(0,2,1)}^{(2,0,1)} [ze^x \, dx + 2yz \, dy + (e^x + y^2) \, dz]$ 20. $\int_{(0,0,0)}^{(1,4,2)} (3y^2 e^x \, dy + y^3 e^x \, dz)$

Chapter 7

Matrices and Determinants. Systems of Linear Equations

Matrix calculus is an elegant and powerful mathematical language in connection with linear equations, linear transformations, systems of differential equations, etc. For this reason matrices have become more and more important in physics, engineering, statistics, and other fields. The development in this direction started about thirty years ago and will continue even more rapidly in the future.

The present chapter will be devoted to matrices (including matrix eigenvalue problems), determinants, and linear equations. In Secs. 7.1 and 7.2 we shall present the basic concepts and rules of matrix algebra. Then we shall consider determinants and systems of linear equations (Secs. 7.3–7.10), and finally matrix eigenvalue problems.

The chapter includes numerical methods for solving systems of linear equations and matrix eigenvalue problems.

Prerequisite for this chapter: vector algebra (Secs. 5.1–5.7).

Sections which may be omitted in a shorter course: 7.9, 7.13–7.15.

References: Appendix 1, Part D.

Answers to problems: Appendix 2.

7.1 Basic Concepts. Addition of Matrices

The concept of a matrix arises in connection with linear relations such as linear transformations and systems of linear equations. For example, consider the linear transformation

$$y_1 = a_{11}x_1 + a_{12}x_2$$

$$y_2 = a_{21}x_1 + a_{22}x_2$$

where the coefficients a_{11} , a_{12} , a_{21} , and a_{22} are given numbers while x_1 , x_2 and also y_1 , y_2 are variable quantities (for instance, the coordinates of two

different coordinate systems in the plane). Arranging the coefficients in the way they occur in the transformation and enclosing them in parentheses, we obtain an array of the form

$$\begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{matrix}$$

This array is an example of a matrix.

Let us now define the concept of a matrix and some related basic notions.

A rectangular array of (real or complex) numbers of the form

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

is called a **matrix**. The numbers a_{11}, \dots, a_{mn} are called the **elements** of the matrix. The horizontal lines are called **rows** or **row vectors**, and the vertical lines are called **columns** or **column vectors** of the matrix. A matrix with m rows and n columns is called an $(m \times n)$ matrix (read "m by n matrix").

Matrices will be denoted by capital (upper case) bold-faced letters **A**, **B**, etc., or by (a_{jk}) , (b_{jk}) , etc., that is, by writing the general element of the matrix, enclosed in parentheses.

In the double-subscript notation for the elements, the first subscript always denotes the row and the second subscript the column containing the given element.

A **matrix**

$$(a_1 \quad a_2 \quad \cdots \quad a_n)$$

having only one row is called a **row matrix** or **row vector**. A matrix

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

having only one column is called a **column matrix** or **column vector**. Row and column matrices will be denoted by small (lower case) bold-faced letters, in agreement with the notation used in vector calculus.

Two $(m \times n)$ matrices $\mathbf{A} = (a_{jk})$ and $\mathbf{B} = (b_{jk})$ are said to be **equal** if, and only if, corresponding elements are equal, that is,

$$a_{jk} = b_{jk} \quad \text{for all } j = 1, \dots, m, \quad k = 1, \dots, n.$$

Then we write

$$\mathbf{A} = \mathbf{B}.$$

Note that this definition of equality refers only to matrices which have the *same number of rows and the same number of columns*.

We shall now consider the addition of matrices.

The sum of two $(m \times n)$ matrices $\mathbf{A} = (a_{jk})$ and $\mathbf{B} = (b_{jk})$ is defined as the $(m \times n)$ matrix $\mathbf{C} = (c_{jk})$ whose elements are

$$(1) \quad c_{jk} = a_{jk} + b_{jk}, \quad \begin{matrix} j = 1, \dots, m, \\ k = 1, \dots, n, \end{matrix}$$

and we write

$$\mathbf{C} = \mathbf{A} + \mathbf{B}.$$

Similarly, the matrix

$$\mathbf{D} = \mathbf{A} - \mathbf{B} = (a_{jk} - b_{jk})$$

is called the *difference of A and B*.

Note that addition and subtraction are defined only for matrices having the same number of rows and the same number of columns.

Example 1. Let

$$\mathbf{A} = \begin{pmatrix} -4 & 6 & 3 \\ 0 & 1 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 5 & -1 & 0 \\ 3 & 1 & 0 \end{pmatrix}.$$

Then

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} 1 & 5 & 3 \\ 3 & 2 & 2 \end{pmatrix}, \quad \mathbf{A} - \mathbf{B} = \begin{pmatrix} -9 & 7 & 3 \\ -3 & 0 & 2 \end{pmatrix}.$$

The product of a matrix $\mathbf{A} = (a_{jk})$ by a number c is defined as the matrix (ca_{jk}) and is denoted by $c\mathbf{A}$ or \mathbf{Ac} ; thus

$$c\mathbf{A} = \mathbf{Ac} = \begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{pmatrix}.$$

Instead of $(-1)\mathbf{A}$ we simply write $-\mathbf{A}$, and $\mathbf{A} + (-\mathbf{B})$ will be written $\mathbf{A} - \mathbf{B}$, in agreement with our definition of subtraction.

Example 2. Let

$$\mathbf{A} = \begin{pmatrix} 2.7 & -1.8 \\ 0.9 & 3.6 \end{pmatrix}.$$

Then

$$\mathbf{A} + \mathbf{A} = 2\mathbf{A} = \begin{pmatrix} 5.4 & -3.6 \\ 1.8 & 7.2 \end{pmatrix}, \quad \frac{10}{9}\mathbf{A} = \begin{pmatrix} 3 & -2 \\ 1 & 4 \end{pmatrix}.$$

The reader may prove that for any $(m \times n)$ matrices \mathbf{A} , \mathbf{B} , \mathbf{C} and any constants a , b , c the following laws hold:

$$(2) \quad \begin{aligned} \mathbf{A} + \mathbf{B} &= \mathbf{B} + \mathbf{A} \\ \mathbf{A} + (\mathbf{B} + \mathbf{C}) &= (\mathbf{A} + \mathbf{B}) + \mathbf{C} \\ c(\mathbf{A} + \mathbf{B}) &= c\mathbf{A} + c\mathbf{B} \\ (a + b)\mathbf{A} &= a\mathbf{A} + b\mathbf{A} \\ a(b\mathbf{A}) &= (ab)\mathbf{A}. \end{aligned}$$

Another frequently used operation is the *transposition of a matrix*. The transpose of an $(m \times n)$ matrix $\mathbf{A} = (a_{jk})$ is defined as the $(n \times m)$ matrix (a_{kj}) and is denoted by \mathbf{A}^T . Thus

$$\mathbf{A}^T = (a_{kj}) = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

and we see that the row vectors of \mathbf{A} become the column vectors of \mathbf{A}^T and vice versa. The reader may prove that

$$(3) \quad (\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T.$$

Example 3. If

$$\mathbf{A} = \begin{pmatrix} 5 & -8 & 1 \\ 4 & 0 & 0 \end{pmatrix}, \quad \text{then} \quad \mathbf{A}^T = \begin{pmatrix} 5 & 4 \\ -8 & 0 \\ 1 & 0 \end{pmatrix}.$$

Example 4. If

$$\mathbf{b} = \begin{pmatrix} 7 & 5 & -2 \end{pmatrix}, \quad \text{then} \quad \mathbf{b}^T = \begin{pmatrix} 7 \\ 5 \\ -2 \end{pmatrix}$$

The elements of a matrix may be real or complex numbers. If all the elements of a matrix are real, the matrix is called a **real matrix**.

A matrix having the same number of rows and columns is called a **square matrix**, and the number of rows is called its **order**. The diagonal containing the elements a_{11} , a_{22} , \dots , a_{nn} of the square matrix is called the **principal diagonal**. Square matrices are of particular importance.

A real square matrix $\mathbf{A} = (a_{jk})$ is said to be **symmetric** if it is equal to its transpose,

$$(4) \quad \mathbf{A}^T = \mathbf{A}, \quad \text{that is,} \quad a_{kj} = a_{jk} \quad (j, k = 1, \dots, n).$$

A real square matrix $\mathbf{A} = (a_{jk})$ is said to be **skew-symmetric** if

$$(5) \quad \mathbf{A}^T = -\mathbf{A}, \quad \text{that is,} \quad a_{kj} = -a_{jk} \quad (j, k = 1, \dots, n).$$

Note that for $k = j$ in (5) we have $a_{jj} = -a_{jj}$, which implies that the elements in the principal diagonal of a skew-symmetric matrix are all zero.

Any real square matrix A may be written as the sum of a symmetric matrix R and a skew-symmetric matrix S , where

$$R = \frac{1}{2}(A + A^T) \quad \text{and} \quad S = \frac{1}{2}(A - A^T).$$

Example 5. The matrices

$$A = \begin{pmatrix} -3 & 1 & 5 \\ 1 & 0 & -2 \\ 5 & -2 & 4 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & -4 & 1 \\ 4 & 0 & -5 \\ -1 & 5 & 0 \end{pmatrix}$$

are symmetric and skew-symmetric, respectively. The matrix

$$A = \begin{pmatrix} 2 & 3 \\ 5 & -1 \end{pmatrix}$$

may be written in the form $A = R + S$ where

$$R = \frac{1}{2}(A + A^T) = \begin{pmatrix} 2 & 4 \\ 4 & -1 \end{pmatrix} \quad \text{and} \quad S = \frac{1}{2}(A - A^T) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

are symmetric and skew-symmetric, respectively.

A square matrix $A = (a_{jk})$ whose elements above the principal diagonal (or below the principal diagonal) are all zero is called a triangular matrix.

For example,

$$T_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 3 & 0 \\ 5 & 0 & 2 \end{pmatrix} \quad \text{and} \quad T_2 = \begin{pmatrix} 1 & 6 & -1 \\ 0 & 2 & 3 \\ 0 & 0 & 4 \end{pmatrix}$$

are triangular matrices.

A square matrix $A = (a_{jk})$ whose elements above and below the principal diagonal are all zero, that is, $a_{jk} = 0$ for all $j \neq k$, is called a diagonal matrix.

For example,

$$\begin{matrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -4 \end{matrix} \quad \text{and} \quad \begin{matrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{matrix}$$

are diagonal matrices.

A diagonal matrix whose elements in the principal diagonal are all 1 is called a unit matrix and will be denoted by I . For example, the three-rowed unit matrix is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

A matrix whose elements are all zero is called a **zero matrix**, or **null matrix**, and is denoted by **0**. Clearly, for any $(m \times n)$ matrix **A**,

$$\mathbf{A} + \mathbf{0} = \mathbf{0} + \mathbf{A} = \mathbf{A},$$

where **0** is the $(m \times n)$ null matrix.

Problems

Let $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} -2 & 0 \\ 3 & 4 \end{pmatrix}$, $\mathbf{C} = \begin{pmatrix} 3 & 2 & 0 \\ 1 & 0 & 4 \end{pmatrix}$. Find

1. $\mathbf{A} + \mathbf{B}$, $\mathbf{B} + \mathbf{A}$
2. $\mathbf{A} - \mathbf{B}$, $\mathbf{B} - \mathbf{A}$
3. $4\mathbf{A} - 2\mathbf{B}$, $2(2\mathbf{A} - \mathbf{B})$
4. \mathbf{A}^T , \mathbf{B}^T , \mathbf{C}^T , $(\mathbf{B}^T)^T$
5. $(\mathbf{A} + \mathbf{B})^T$, $\mathbf{A}^T + \mathbf{B}^T$
6. $\mathbf{A} + \mathbf{A}^T$, $\mathbf{A} - \mathbf{A}^T$
7. Is $\mathbf{C} + \mathbf{C}^T$ defined?
8. Is $\mathbf{A} + \mathbf{C}$ defined?
9. Are \mathbf{A} and \mathbf{B} symmetric?
10. Represent \mathbf{B} as the sum of a symmetric and a skew-symmetric matrix.

Let $\mathbf{M} = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 3 & -4 & 0 \end{pmatrix}$, $\mathbf{N} = \begin{pmatrix} 3 & 8 \\ 4 & -1 \\ 7 & 2 \end{pmatrix}$, $\mathbf{P} = \begin{pmatrix} 2 & 1 & 4 \\ 8 & 0 & 0 \\ 3 & -1 & 1 \end{pmatrix}$. Find

11. $\mathbf{M} + \mathbf{P}$, $\mathbf{P} + \mathbf{M}$
12. $\mathbf{M} - \mathbf{P}$, $\mathbf{P} - \mathbf{M}$
13. $3\mathbf{M} + 2\mathbf{P}$
14. $\mathbf{M} + \mathbf{M}^T$, $\mathbf{M} - \mathbf{M}^T$
15. $(\mathbf{M} - \mathbf{P})^T$, $\mathbf{M}^T - \mathbf{P}^T$
16. $(3\mathbf{P})^T$, $3\mathbf{P}^T$
17. Are $\mathbf{M} + \mathbf{N}$, $\mathbf{M} + \mathbf{N}^T$, $\mathbf{M} + \mathbf{P}^T$, $\mathbf{N} + \mathbf{C}^T$ (see before) defined?

Prove

18. The formulas (2)
19. $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
20. $(a\mathbf{A})^T = a\mathbf{A}^T$, $(\mathbf{A}^T)^T = \mathbf{A}$

7.2 Matrix Multiplication

We shall now define the multiplication of a matrix by a matrix. This definition will be suggested by the use of matrices in connection with linear transformations. To illustrate this, let us start with a very simple case.

We consider three coordinate systems in the plane which we denote as the x_1x_2 -system, the y_1y_2 -system, and the z_1z_2 -system, and we assume that these systems are related by the linear transformations

$$(1) \quad \begin{aligned} x_1 &= a_{11}y_1 + a_{12}y_2 \\ x_2 &= a_{21}y_1 + a_{22}y_2 \end{aligned}$$

and

$$(2) \quad \begin{aligned} y_1 &= b_{11}z_1 + b_{12}z_2 \\ y_2 &= b_{21}z_1 + b_{22}z_2. \end{aligned}$$

Then it is clear that the x_1x_2 -coordinates can be obtained directly from the z_1z_2 -coordinates by a single linear transformation of the form

$$(3) \quad \begin{aligned} x_1 &= c_{11}z_1 + c_{12}z_2 \\ x_2 &= c_{21}z_1 + c_{22}z_2 \end{aligned}$$

whose coefficients can be found by inserting (2) into (1),

$$\begin{aligned}x_1 &= a_{11}(b_{11}z_1 + b_{12}z_2) + a_{12}(b_{21}z_1 + b_{22}z_2) \\x_2 &= a_{21}(b_{11}z_1 + b_{12}z_2) + a_{22}(b_{21}z_1 + b_{22}z_2)\end{aligned}$$

and comparing this and (3); the result is

$$\begin{aligned}c_{11} &= a_{11}b_{11} + a_{12}b_{21} & c_{12} &= a_{11}b_{12} + a_{12}b_{22} \\c_{21} &= a_{21}b_{11} + a_{22}b_{21} & c_{22} &= a_{21}b_{12} + a_{22}b_{22}\end{aligned}$$

or briefly

$$(4) \quad c_{jk} = a_{j1}b_{1k} + a_{j2}b_{2k} = \sum_{i=1}^2 a_{ji}b_{ik} \quad j, k = 1, 2.$$

Now the coefficient matrices of the transformations (1) and (2) are

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

and we define the product \mathbf{AB} (in this order) of \mathbf{A} and \mathbf{B} to be the coefficient matrix

$$\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

of the "composed transformation" (3), that is,

$$\mathbf{C} = \mathbf{AB}$$

where the elements of \mathbf{C} are given by (4).

If we write \mathbf{A} in terms of row vectors, say,

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{pmatrix} \quad \text{where} \quad \begin{aligned} \mathbf{a}_1 &= (a_{11} \quad a_{12}) \\ \mathbf{a}_2 &= (a_{21} \quad a_{22}) \end{aligned}$$

and \mathbf{B} in terms of column vectors, say,

$$\mathbf{B} = (\mathbf{b}_1 \quad \mathbf{b}_2) \quad \text{where} \quad \mathbf{b}_1 = \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} b_{12} \\ b_{22} \end{pmatrix}$$

then the elements of \mathbf{C} may simply be written as scalar products:

$$\begin{aligned}c_{11} &= \mathbf{a}_1 \cdot \mathbf{b}_1, & c_{12} &= \mathbf{a}_1 \cdot \mathbf{b}_2, \\c_{21} &= \mathbf{a}_2 \cdot \mathbf{b}_1, & c_{22} &= \mathbf{a}_2 \cdot \mathbf{b}_2.\end{aligned}$$

Example 1. Let

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 1 & -2 \\ 5 & 3 \end{pmatrix}.$$

Then, by (4),

$$\mathbf{AB} = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 5 & 3 \end{pmatrix} = \begin{pmatrix} 2 \cdot 1 + 1 \cdot 5 & 2 \cdot (-2) + 1 \cdot 3 \\ 3 \cdot 1 + 4 \cdot 5 & 3 \cdot (-2) + 4 \cdot 3 \end{pmatrix} = \begin{pmatrix} 7 & -1 \\ 23 & 6 \end{pmatrix}.$$

To arrive at the definition of the multiplication of two matrices in the general case, we now extend the preceding consideration to two general linear transformations, say, the transformation

$$(5) \quad \begin{aligned} x_1 &= a_{11}y_1 + \cdots + a_{1n}y_n \\ x_2 &= a_{21}y_1 + \cdots + a_{2n}y_n \\ &\dots \\ x_m &= a_{m1}y_1 + \cdots + a_{mn}y_n \end{aligned}$$

and the transformation

$$(6) \quad \begin{aligned} y_1 &= b_{11}z_1 + \cdots + b_{1p}z_p \\ y_2 &= b_{21}z_1 + \cdots + b_{2p}z_p \\ &\dots \\ y_n &= b_{n1}z_1 + \cdots + b_{np}z_p \end{aligned}$$

By inserting (6) into (5) we find that x_1, \dots, x_m can be expressed in terms of z_1, \dots, z_p in the form

$$(7) \quad \begin{aligned} x_1 &= c_{11}z_1 + \cdots + c_{1p}z_p \\ x_2 &= c_{21}z_1 + \cdots + c_{2p}z_p \\ &\dots \\ x_m &= c_{m1}z_1 + \cdots + c_{mp}z_p \end{aligned}$$

where the coefficients are given by the formula

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk}$$

The requirement that the product AB (in this order) of the coefficient matrices $A = (a_{jk})$ and $B = (b_{jk})$ of the transformations (5) and (6) be the coefficient matrix $C = (c_{jk})$ of the "composed transformation" (7) leads to the following definition.

Definition. Let $\mathbf{A} = (a_{jk})$ be an $(m \times n)$ matrix and $\mathbf{B} = (b_{jk})$ an $(r \times p)$ matrix, then the product \mathbf{AB} (in this order) is defined only when $r = n$ and is the $(m \times p)$ matrix $\mathbf{C} = (c_{jk})$ whose elements are

$$(8) \quad c_{jk} = a_{j1}b_{1k} + a_{j2}b_{2k} + \cdots + a_{jn}b_{nk} = \sum_{i=1}^n a_{ji}b_{ik}.$$

Instead of \mathbf{AA} we simply write \mathbf{A}^2 , etc.

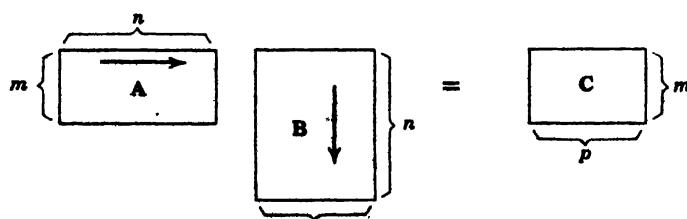


Fig. 201. Matrix multiplication. $AB = C$

We see that c_{jk} is the scalar product of the j -th row vector of the first matrix, \mathbf{A} , and the k -th column vector of the second matrix, \mathbf{B} ; that is, if we write

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_m \end{pmatrix} \quad \text{and} \quad \mathbf{B} = (\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_p)$$

where $\mathbf{a}_1, \dots, \mathbf{a}_m$ are the row vectors of \mathbf{A} and $\mathbf{b}_1, \dots, \mathbf{b}_p$ are the column vectors of \mathbf{B} , then

$$(9) \quad \mathbf{C} = \begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 & \cdots & \mathbf{a}_1 \cdot \mathbf{b}_p \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 & \cdots & \mathbf{a}_2 \cdot \mathbf{b}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_m \cdot \mathbf{b}_1 & \mathbf{a}_m \cdot \mathbf{b}_2 & \cdots & \mathbf{a}_m \cdot \mathbf{b}_p \end{pmatrix}.$$

The process of matrix multiplication is, therefore, conveniently referred to as the **multiplication of rows into columns**. Cf. Fig. 201 on p. 363.

Example 2. Let

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & -1 \\ 0 & 4 & 6 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 1 & 0 & 2 \\ 5 & 3 & 1 \\ 6 & 4 & 2 \end{pmatrix}.$$

Then, by (8), the product \mathbf{AB} is

$$\mathbf{AB} = \begin{pmatrix} 3 & 2 & -1 \\ 0 & 4 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 5 & 3 & 1 \\ 6 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 7 & 2 & 6 \\ 56 & 36 & 16 \end{pmatrix},$$

whereas the product \mathbf{BA} is not defined.

The multiplication of matrices by vectors and the matrix multiplication of vectors may be illustrated by the following examples.

Example 3.

$$\begin{pmatrix} 3 & 4 & 2 \\ 6 & 0 & -1 \\ -5 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 19 \\ 4 \\ -9 \end{pmatrix}$$

Example 4.

$$(3 \quad 6 \quad 1) \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} = (19), \quad \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} (3 \quad 6 \quad 1) = \begin{pmatrix} 3 & 6 & 1 \\ 6 & 12 & 2 \\ 12 & 24 & 4 \end{pmatrix}$$

Note that the first of these products corresponds to the scalar product as defined in vector algebra.

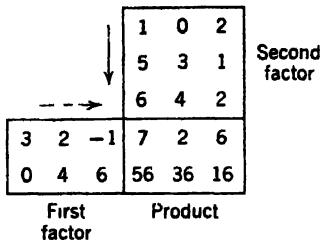


Fig. 202. Matrix multiplication.

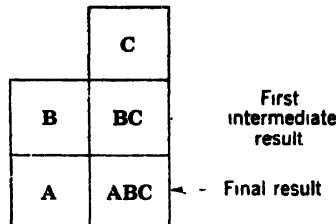


Fig. 203. Matrix product ABC.

Numerical work can be checked by the use of the sums of the elements of the columns. In Ex. 2,

$$\begin{pmatrix} 3 & 2 & -1 \\ 0 & 4 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 5 & 3 & 1 \\ 6 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 7 & 2 & 6 \\ 56 & 36 & 16 \end{pmatrix}$$

Sum 3 6 5 63 38 22

Check:

$$3 \cdot 1 + 6 \cdot 5 + 5 \cdot 6 = 63.$$

$$3 \cdot 0 + 6 \cdot 3 + 5 \cdot 4 = 38$$

$$3 \cdot 2 + 6 \cdot 1 + 5 \cdot 2 = 22$$

Figure 202 shows an arrangement of the matrices A and B in Ex. 2 and their product AB , which is very convenient for numerical work. The point is that each element c_{jk} of the product matrix occupies the intersection of that row of the first matrix and that column of the last matrix which are used for computing c_{jk} . The extension of the arrangement to computations of products of three and more matrices is obvious (Fig. 203).

By using the definition of matrix multiplication we may now write the linear transformation (5) in the simple form

$$(5') \quad \mathbf{x} = \mathbf{A}\mathbf{y}$$

where $\mathbf{A} = (a_{jk})$ and

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Similarly, (6) may be written

$$(6') \quad \mathbf{y} = \mathbf{B}\mathbf{z},$$

where \mathbf{z} is the column matrix with elements z_1, \dots, z_n . By inserting (6') into (5') and using the associative law of matrix multiplication (see right below) we obtain

$$(7') \quad \mathbf{x} = \mathbf{A}(\mathbf{B}\mathbf{z}) = \mathbf{ABx} = \mathbf{Cz},$$

in agreement with (7).

Matrix multiplication is associative and distributive; that is,

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}) = \mathbf{ABC}$$

$$(10) \quad (\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$$

$$\mathbf{C}(\mathbf{A} + \mathbf{B}) = \mathbf{CA} + \mathbf{CB}$$

provided \mathbf{A} , \mathbf{B} , and \mathbf{C} are such that the expressions on the left are defined. The proofs are left as an exercise.

The following two properties of matrix multiplication have no counterparts in the usual multiplication of numbers and should, therefore, be carefully observed.

Matrix multiplication is not commutative; that is, if \mathbf{A} and \mathbf{B} are matrices such that both \mathbf{AB} and \mathbf{BA} are defined, then

$$\mathbf{AB} \neq \mathbf{BA} \quad \text{in general.}$$

For example,

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{but} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Cf. also Ex. 4, above.

It is clear that *the order of the factors in a matrix product must be carefully observed*, and only in certain cases the equation $\mathbf{AB} = \mathbf{BA}$ will hold. To be precise we say that, in the product \mathbf{AB} , the matrix \mathbf{B} is **premultiplied** by the matrix \mathbf{A} , or, alternatively, that \mathbf{A} is **postmultiplied** by \mathbf{B} .

The other unusual property of matrix multiplication can be seen from the example

$$\mathbf{AB} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{where} \quad \mathbf{A} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

which illustrates that

$$\mathbf{AB} = \mathbf{0} \quad \text{does not imply} \quad \mathbf{A} = \mathbf{0} \text{ or } \mathbf{B} = \mathbf{0}.$$

We shall consider this surprising fact in more detail in Sec. 7.6. The reader may show that in the example, $\mathbf{BA} \neq \mathbf{0}$.

A diagonal matrix whose diagonal elements are all equal is called a scalar matrix. If

$$S = \begin{pmatrix} k & 0 & \cdots & 0 \\ 0 & k & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & k \end{pmatrix}$$

is an n -rowed scalar matrix and A is any n -rowed square matrix, then

$$(11) \quad AS = SA = kA;$$

that is, the multiplication of an n -rowed square matrix A by an n -rowed scalar matrix has the same effect as the multiplication of A by a number, and S commutes with any n -rowed square matrix. The simple proof is left to the reader.

The transpose of a product equals the product of the transposed factors, taken in reverse order,

$$(12) \quad (AB)^T = B^T A^T.$$

The proof follows from the definition of matrix multiplication and is left to the reader.

If b_1, \dots, b_s are the column vectors of a matrix B and the product AB is defined, then from the definition of matrix multiplication it follows immediately that the product matrix AB has the column vectors Ab_1, \dots, Ab_s ,

$$(13) \quad AB = (Ab_1 \ Ab_2 \ \cdots \ Ab_s).$$

Problems

Let

$$A = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \\ 1 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 3 \\ 4 & 0 & -1 \end{pmatrix}$$

Find those of the following expressions which are defined.

- | | | |
|------------------------------------|------------------------------------|--------------------------|
| 1. AB , $A^T B$, BA , $B^T A$ | 2. BC , $B^T C$, CB , $C^T B$ | 3. A^2 , B^2 , C^2 |
| 4. AAT , A^TA | 5. BB^T , B^TB | 6. CC^T , C^TC |
| 7. C^2 , C^4 | 8. BCA , $B^T CA$ | 9. $A^T CA$ |

10. The so-called Pauli spin matrices are

$$S_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad S_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (i = \sqrt{-1}).$$

Show that $S_x S_y = i S_z$, $S_y S_z = -i S_x$, $S_x^2 = S_y^2 = S_z^2 = I_2$.

Find \mathbf{AB} and \mathbf{BA} where \mathbf{A} and \mathbf{B} are

✓ 11. $\begin{pmatrix} 2 & -1 & 0 \\ 0 & -2 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \begin{pmatrix} -2 & 1 & -1 \\ 1 & 2 & -2 \\ 2 & -1 & -4 \end{pmatrix}$

12. $\begin{pmatrix} 1 & -2 & 4 \\ 3 & 1 & 5 \\ 2 & 4 & 0 \end{pmatrix}, \begin{pmatrix} 2 & 4 & -2 \\ -1 & -2 & 1 \\ -1 & -2 & 1 \end{pmatrix}$

13. Let $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, $\mathbf{A} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$. Show that the linear transformation $\mathbf{y} = \mathbf{Ax}$ represents a rotation of the Cartesian x_1x_2 -coordinate system in the plane about the origin.

14. Interpret the transformation $\mathbf{x} = \mathbf{Bz}$ geometrically, where

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

15. Consider the transformations $\mathbf{y} = \mathbf{Ax}$ and $\mathbf{x} = \mathbf{Bz}$ in Probs. 13 and 14. Find the composite transformation which expresses \mathbf{y} in terms of \mathbf{z} . What does it mean geometrically that $\mathbf{AB} \neq \mathbf{BA}$?

16. Represent each of the transformations

$$\begin{aligned} x_1 &= 2y_1 - y_2 & y_1 &= z_1 + z_2 \\ x_2 &= -y_1 + 2y_2 & y_2 &= z_1 - z_2 \end{aligned}$$

by the use of matrices and find the composite transformation which expresses x_1, x_2 in terms of z_1, z_2 .

17. Same task as in Prob. 16 for the transformations

$$\begin{aligned} x_1 &= y_1 + 2y_2 & y_1 &= 2z_2 \\ x_2 &= y_1 - y_2 & y_2 &= z_1 + z_2 \end{aligned}$$

Prove

18. The formula (11)

19. $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$

20. $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$

7.3 Determinants

Determinants arise in connection with simultaneous linear algebraic equations. The solutions of the equations are rather unmanageable when written at length, but become simple-looking expressions by using determinants.

For example, the system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= k_1 \\ a_{21}x_1 + a_{22}x_2 &= k_2 \end{aligned}$$

has the solution

$$x_1 = \frac{k_1 a_{22} - k_2 a_{12}}{a_{11} a_{22} - a_{12} a_{21}}, \quad x_2 = \frac{a_{11} k_2 - a_{21} k_1}{a_{11} a_{22} - a_{12} a_{21}},$$

provided $a_{11}a_{22} - a_{12}a_{21}$ is not zero (cf. Sec. 0.3). This expression in the denominator is written in the form

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

and is called a **determinant of second order**; thus

$$(1) \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

The solution may then be written

$$x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}$$

where

$$D = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad D_1 = \begin{vmatrix} k_1 & a_{12} \\ k_2 & a_{22} \end{vmatrix}, \quad D_2 = \begin{vmatrix} a_{11} & k_1 \\ a_{21} & k_2 \end{vmatrix}$$

Systems of equations in three unknowns suggest the introduction of "third-order determinants," and so on. We shall now define determinants of order n and consider their most important properties. The use of determinants in connection with systems of linear equations will be explained in Sec. 7.5.

A determinant of order n is a square array of n^2 quantities enclosed between vertical bars,

$$(2) \quad D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

which has a certain value defined below. The quantities a_{11}, \dots, a_{nn} which are numbers (or sometimes functions), are called the **elements** of the determinant. The horizontal lines of elements are called **rows**; the vertical lines are called **columns**. The sloping line of elements extending from a_{11} to a_{nn} is called the **principal diagonal** of the determinant.

By deleting the i th row and the k th column from the determinant D we obtain an $(n-1)$ th order determinant (a square array of $n-1$ rows and $n-1$ columns between vertical bars), which is called the **minor** of the element a_{ik} (which belongs to the deleted row and column) and is denoted by M_{ik} .

The minor M_{ik} multiplied by $(-1)^{i+k}$ is called the **cofactor** of a_{ik} and will be denoted by C_{ik} ; thus

$$(3) \quad C_{ik} = (-1)^{i+k} M_{ik}.$$

For example, in the third-order determinant

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

we have

$$C_{11} = M_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \quad C_{32} = -M_{32} = -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}, \text{ etc.}$$

We are now in a position to define an n th-order determinant in a way which is most convenient for practical purposes.

Definition. *The symbol*

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

is called a determinant of order n and means the sum of the products of the elements of any row or column and their respective cofactors; that is,

$$(4a) \quad D = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \quad (i = 1, 2, \dots, \text{or } n)$$

or

$$(4b) \quad D = a_{1k}C_{1k} + a_{2k}C_{2k} + \cdots + a_{nk}C_{nk} \quad (k = 1, 2, \dots, \text{or } n).$$

In this way, D is defined in terms of n determinants of order $n - 1$, each of which is, in turn, defined in terms of $n - 1$ determinants of order $n - 2$, and so on; we finally arrive at second-order determinants, in which the cofactors of the elements are single elements of D . We see that if the elements of D are numbers, the value of D will be a number.

Furthermore, it follows from the definition that we may develop D by any row or column, that is, choose in (4) the elements of any row or column, similarly when developing the cofactors in (4), and so on. Consequently, before we accept this definition we have to show that it is unambiguous, that is, yields the same value of D no matter which columns or rows we choose.

For a second-order determinant

$$D = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

this is immediately clear, because then we have only four possible forms of (4), namely, the development by

$$\text{the first row: } D = a_{11}a_{22} + a_{12}(-a_{21}),$$

$$\text{the last row: } D = a_{21}(-a_{12}) + a_{22}a_{11},$$

$$\text{the first column: } D = a_{11}a_{22} + a_{21}(-a_{12}),$$

$$\text{and the last column: } D = a_{12}(-a_{21}) + a_{22}a_{11}.$$

We see that we always obtain the same value of D , which value agrees with (1).

We shall now prove that *our definition of a determinant D of arbitrary order n yields the same value of D no matter which row or column is chosen.*

We shall prove first that *the same result is obtained no matter which row is chosen.*

The proof proceeds by induction. The statement is true for a second order determinant (see before). Assuming that it is true for a determinant of order $n - 1$, we prove that it is true for a determinant D of order n .

For this purpose we expand D , given in the definition, in terms of each of two arbitrary rows, say the i th and the j th, and compare the results. Without loss of generality let us assume $i < j$.

First expansion. We expand D by the i th row. A typical term in this expansion is

$$(5) \quad a_{ik}C_{ik} = a_{ik} \cdot (-1)^{i+k}M_{ik}.$$

The minor M_{ik} of a_{ik} in D is an $(n - 1)$ th order determinant. By the induction hypothesis we may expand it by any row. We expand it by the row corresponding to the j th row of D . This row contains the elements a_{jl} ($l \neq k$). It is the $(j - 1)$ th row of M_{ik} , because M_{ik} does not contain elements of the i th row of D , and $i < j$. We have to distinguish between two cases as follows.

Case I. If $l < k$, then the element a_{jl} belongs to the l th column of M_{ik} (cf. Fig. 204 on p. 372). Hence the term involving a_{jl} in this expansion is

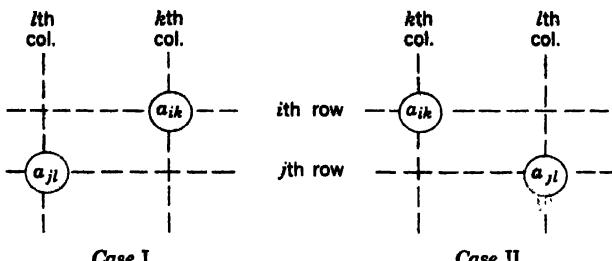
$$(6) \quad a_{jl} \cdot (\text{cofactor of } a_{jl} \text{ in } M_{ik}) = a_{jl} \cdot (-1)^{(j-1)+l}M_{ikjl}$$

where M_{ikjl} is the minor of a_{jl} in M_{ik} . Since this minor is obtained from M_{ik} by deleting the row and column of a_{jl} , it is obtained from D by deleting the i th and j th rows and the k th and l th columns of D . We insert the expansions of the M_{ik} into that of D . Then it follows from (5) and (6) that the terms of the resulting representation of D are of the form

$$(71) \quad a_{ik}a_{jl} \cdot (-1)^b M_{ikjl} \quad (l < k)$$

where

$$b = i + k + j + l - 1. \quad ?$$

Fig. 204. Cases I and II of the two expansions of D .

Case II. If $l > k$, the only difference is that then a_{jl} belongs to the $(l - 1)$ th column of M_{ik} , because M_{ik} does not contain elements of the k th column of D , and $k < l$. This causes an additional minus sign in (6), and, instead of (7I), we therefore obtain

$$(7II) \quad -a_{ik}a_{jl} \cdot (-1)^b M_{ikjl} \quad (l > k)$$

where b is the same as before.

Second expansion. We now expand D at first by the j th row. A typical term in this expansion is

$$(8) \quad a_{jl}C_{jl} = a_{jl} \cdot (-1)^{j+i} M_{jl}.$$

By induction hypothesis we may expand the minor M_{jl} of a_{jl} in D by its i th row, which corresponds to the i th row of D , since $j > i$.

Case I. If $k > l$, the element a_{ik} in that row belongs to the $(k - 1)$ th column of M_{jl} , because M_{jl} does not contain elements of the l th column of D , and $l < k$ (cf. Fig. 204). Hence the term involving a_{ik} in this expansion is

$$(9) \quad a_{ik} \cdot (\text{cofactor of } a_{ik} \text{ in } M_{jl}) = a_{ik} \cdot (-1)^{i+(k-1)} M_{ikjl},$$

where the minor M_{ikjl} of a_{ik} in M_{jl} is obtained by deleting the i th and j th rows and the k th and l th columns of D [and is, therefore, identical with M_{ikjl} in (6), so that our notation is consistent]. We insert the expansions of the M_{jl} into that of D . It follows from (8) and (9) that this yields a representation whose terms are identical with those given by (7I) when $l < k$.

Case II. If $k < l$, then a_{ik} belongs to the k th column of M_{jl} , we obtain an additional minus sign, and the result agrees with that characterized by (7II).

We have shown that the two expansions of D consist of the same terms, and this proves our statement concerning rows.

The proof of the statement concerning columns is quite similar; if we expand D in terms of two arbitrary columns, say, the k th and the l th, we find that the general term involving $a_{jl}a_{ik}$ is exactly the same as before. This proves that not only all column expansions of D yield the same value, but also that their common value is equal to the common value of the row expansions of D .

This completes the proof and shows that our definition of an n th order determinant is unambiguous.

Example 1. Let

$$D = \begin{vmatrix} 1 & 3 & 0 \\ 2 & 6 & 4 \\ -1 & 0 & 2 \end{vmatrix}$$

The development by the first row is

$$D = 1 \begin{vmatrix} 6 & 4 \\ 0 & 2 \end{vmatrix} - 3 \begin{vmatrix} 2 & 4 \\ -1 & 2 \end{vmatrix} = 1(12 - 0) - 3(4 + 4) = -12.$$

The development by the first column is

$$D = 1 \begin{vmatrix} 6 & 4 \\ 0 & 2 \end{vmatrix} - 2 \begin{vmatrix} 3 & 0 \\ 0 & 2 \end{vmatrix} - 1 \begin{vmatrix} 3 & 0 \\ 6 & 4 \end{vmatrix} = 12 - 12 - 12 = -12,$$

etc.

From our definition we may now readily obtain the most important properties of determinants, as follows.

Since the same value is obtained whether we expand a determinant by any row or any column we have

Theorem 1. *The value of a determinant is not altered if its rows are written as columns, in the same order.*

Example 2.

$$\begin{vmatrix} 1 & 3 & 0 \\ 2 & 6 & 4 \\ -1 & 0 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 \\ 3 & 6 & 0 \\ 0 & 4 & 2 \end{vmatrix} = -12$$

Theorem 2. *If all the elements of one row (or one column) of a determinant are multiplied by the same factor k , the value of the new determinant is k times the value of the given determinant.*

Proof. Expand the determinant by that row (or column) whose elements are multiplied by k .

Note that kD equals a determinant obtained by multiplying the elements of just one row (or column) of D by k , whereas in the case of a matrix, kA equals the matrix obtained by multiplying all the elements of A by k .

Theorem 2 can be used for simplifying a given determinant: cf. the following example.

Example 3.

$$\begin{vmatrix} 1 & 3 & 0 \\ 2 & 6 & 4 \\ -1 & 0 & 2 \end{vmatrix} = 2 \begin{vmatrix} 1 & 3 & 0 \\ 1 & 3 & 2 \\ -1 & 0 & 2 \end{vmatrix} = 6 \begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 2 \\ -1 & 0 & 2 \end{vmatrix} = 12 \begin{vmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ -1 & 0 & 1 \end{vmatrix} = -12$$

From Theorem 2, with $k = 0$, or directly, by expanding, we obtain

Theorem 3. If all the elements of a row (or a column) of a determinant are zero, the value of the determinant is zero.

Theorem 4. If each element of a row (or a column) of a determinant is expressed as a binomial, the determinant can be written as the sum of two determinants, for example,

$$\begin{vmatrix} a_1 + d_1 & b_1 & c_1 \\ a_2 + d_2 & b_2 & c_2 \\ a_3 + d_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}$$

Proof. Expand the determinant by the row (or column) whose terms are binomials.

The generalization of Theorem 4 to the case of elements which are sums of more than two terms is obvious.

Example 4.

$$\begin{vmatrix} 4x+3 & 2 & 1 \\ x & 3 & 4 \\ 2x-1 & 1 & -1 \end{vmatrix} = \begin{vmatrix} 4x & 2 & 1 \\ x & 3 & 4 \\ 2x & 1 & -1 \end{vmatrix} + \begin{vmatrix} 3 & 2 & 1 \\ 0 & 3 & 4 \\ -1 & 1 & -1 \end{vmatrix}$$

Theorem 5. If any two rows (or two columns) of a determinant are interchanged, the value of the determinant is multiplied by -1 .

Proof. Consider the determinant D in the definition. Carry out the development of D whose terms are of the form (7I) and (7II) ("First expansion," p. 371). Let D^* denote the determinant obtained from D by interchanging the i th and j th rows of D . Then the i th row of D^* consists of the elements a_{ji} , $l = 1, \dots, n$, and the j th row consists of the elements a_{ik} , $k = 1, \dots, n$. Develop D^* by its i th row and then each occurring $(n - 1)$ th order determinant by its row corresponding to the j th row of D^* . The terms of the resulting representation of D^* will be quite similar to those of D , given by (7I) and (7II). However, there is one essential difference. The element a_{ji} now appears already in the first step of the development. Hence, if $k > l$, the element a_{ik} appearing when developing the minor of a_{ji} in D^* , will belong to the $(k - 1)$ th column of that minor, because the minor does not contain elements of the l th column of D^* , and $k > l$. This causes an additional minus sign in the formula for D^* corresponding to the formula (6) for D , and this minus sign will appear in the formula corresponding to (7I). Otherwise there will be no difference between these two formulas, because they will both contain $a_{ik}a_{ji}$, and $M_{ikj;i}$ will be as before, since it does not contain elements from the i th and j th rows of D and D^* , the only two rows by which these determinants differ from each other. Similarly, if $l > k$, an additional minus sign will appear in the formula corresponding to (7II). This completes the proof.

Example 5.

$$\begin{vmatrix} 2 & 6 & 4 \\ 1 & 3 & 0 \\ 1 & 0 & 2 \end{vmatrix} = - \begin{vmatrix} 1 & 3 & 0 \\ 2 & 6 & 4 \\ -1 & 0 & 2 \end{vmatrix} = 12$$

Theorem 6. If corresponding elements of two rows (or two columns) of a determinant are proportional, the value of the determinant is zero.

Proof. Let the elements of the i th and j th rows of D be proportional, say, $a_{ik} = ca_{jk}$, $k = 1, \dots, n$. If $c = 0$, then $D = 0$. Let $c \neq 0$. By Theorem 2,

$$D = cB$$

where the i th and j th rows of B are identical. Interchange these rows. Then, by Theorem 5, B goes over into $-B$. On the other hand, since the rows are identical, the new determinant is still B . Thus $B = -B$, $B = 0$, and $D = 0$.

Example 6.

$$\begin{vmatrix} 3 & 6 & -4 \\ 1 & -1 & 3 \\ -6 & -12 & 8 \end{vmatrix} = 0$$

Before evaluating a determinant it is advisable to simplify it. This may be done by Theorem 2 and

Theorem 7. The value of a determinant is left unchanged, if the elements of a row (or column) are altered by adding to them any constant multiple of the corresponding elements in any other row (or column).

Proof. Apply Theorem 4 to the determinant that results from the given addition. This yields a sum of two determinants; one is the original determinant and the other contains two proportional rows. According to Theorem 6, the second determinant is zero and the proof is complete.

Example 7. Evaluate

$$D = \begin{vmatrix} 1 & 24 & 21 & 93 \\ 2 & -37 & -1 & 194 \\ -2 & 35 & 0 & -171 \\ -3 & 177 & 63 & 234 \end{vmatrix}.$$

Add the second row to the third, add three times the first row to the last, subtract two times the first row from the second, and develop the resulting determinant by the first column:

$$D = \begin{vmatrix} 1 & 24 & 21 & 93 \\ 0 & -85 & -43 & 8 \\ 0 & -2 & -1 & 23 \\ 0 & 249 & 126 & 513 \end{vmatrix} = \begin{vmatrix} -85 & -43 & 8 \\ -2 & -1 & 23 \\ 249 & 126 & 513 \end{vmatrix}.$$

Add three times the first row to the last row:

$$\begin{array}{c|ccc} & -85 & -43 & 8 \\ D & \begin{array}{ccc} -2 & -1 & 23 \\ -6 & -3 & 537 \end{array} \end{array}$$

Subtract two times the second column from the first, and then develop the resulting determinant by the first column:

$$D = \begin{vmatrix} 1 & -43 & 8 \\ 0 & -1 & 23 \\ 0 & -3 & 537 \end{vmatrix} = \begin{vmatrix} -1 & 23 \\ -3 & 537 \end{vmatrix} = -537 + 69 = -468.$$

We shall now consider the multiplication of determinants. Of course we may obtain the product of any two determinants by evaluating the determinants and multiplying the results. However, if the determinants are of the same order n , it is often desirable to write their product again as an n th order determinant. This can be done by means of

Theorem 8.

$$(10) \quad \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \begin{vmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \cdots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{vmatrix} = \begin{vmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \cdots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{vmatrix},$$

where the element c_{ik} in the i -th row and k -th column is

$$(10') \quad c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk},$$

the scalar product of the i -th row vector of the first determinant and the k -th column vector of the second determinant.

Remark. Note that this multiplication is performed in the same way as the multiplication of matrices (Sec. 7.2).

Proof of Theorem 8 for $n = 2$. Let

$$D_1 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad D_2 = \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}, \quad \text{and } D = D_1 D_2.$$

We first show that

$$(11) \quad D = D_1 D_2 = \begin{vmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ -1 & 0 & b_{11} & b_{12} \\ 0 & -1 & b_{21} & b_{22} \end{vmatrix}.$$

To prove (11), develop the determinant on the right by the first row and then the resulting two determinants again by their first rows:

$$D = a_{11} \begin{vmatrix} a_{22} & 0 & 0 \\ 0 & b_{11} & b_{12} \\ -1 & b_{21} & b_{22} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & 0 & 0 \\ -1 & b_{11} & b_{12} \\ 0 & b_{21} & b_{22} \end{vmatrix}$$

$$= a_{11}a_{22} \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} - a_{12}a_{21} \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} = (a_{11}a_{22} - a_{12}a_{21})D_2 = D_1D_2.$$

We now transform the determinant in (11) as follows. We add the third row, multiplied by a_{11} , and the fourth row, multiplied by a_{12} , to the first row; then the first row takes the form

$$0 \quad 0 \quad a_{11}b_{11} + a_{12}b_{21} \quad a_{11}b_{12} + a_{12}b_{22}.$$

We now add the third row, multiplied by a_{21} , and the fourth row, multiplied by a_{22} , to the second row; then the second row becomes

$$0 \quad 0 \quad a_{21}b_{11} + a_{22}b_{21} \quad a_{21}b_{12} + a_{22}b_{22}.$$

Altogether,

$$D = \begin{vmatrix} 0 & 0 & a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ 0 & 0 & a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \\ -1 & 0 & b_{11} & b_{12} \\ 0 & -1 & b_{21} & b_{22} \end{vmatrix}.$$

Developing this determinant by the first column, and then the resulting third-order determinant again by its first column, we obtain

$$D = D_1D_2 = \begin{vmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{vmatrix},$$

which is (10) with $n = 2$.

Proof of Theorem 8 for arbitrary n . The steps are quite similar to those in the case $n = 2$. We first show that the product of the given determinants can be written as the $2n$ th-order determinant

$$D = \begin{vmatrix} a_{11} & \cdot & \cdots & a_{1n} & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ a_{n1} & \cdot & \cdots & a_{nn} & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 & b_{11} & \cdots & b_{1n} \\ 0 & -1 & \cdots & 0 & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & & & -1 & & b_{n1} \end{vmatrix}.$$

To prove this, develop the determinant by the first row, the resulting determinants of order $2n - 1$ again by their first rows, etc. After n steps, the result will be of the form

$$(\dots) \begin{vmatrix} b_{11} & \cdots & b_{1n} \\ \cdot & \cdots & \cdot \\ b_{n1} & \cdots & b_{nn} \end{vmatrix}$$

where (\dots) is the representation of the first of the given determinants in terms of products of its elements. Then D is transformed as follows. Add to the first row

the $(n + 1)$ th row multiplied by a_{11} ,
 the $(n + 2)$ th row multiplied by a_{12} ,

 the $2n$ th row multiplied by a_{1n} .

The first row will then become

$$0 \cdots 0 \quad \underbrace{a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1}}_{(n+1)\text{th element} = c_{11}} \cdots \quad \underbrace{a_{11}b_{1n} + \cdots + a_{1n}b_{nn}}_{\text{last element} = c_{1n}}$$

Transforming the second, third, \dots , n th row in a similar fashion, we altogether obtain

$$D = \begin{vmatrix} 0 & \cdots & 0 & c_{11} & \cdots & c_{1n} \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & \cdots & 0 & c_{n1} & \cdots & c_{nn} \\ -1 & \cdots & 0 & b_{11} & \cdots & b_{1n} \\ \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & \cdots & -1 & b_{n1} & \cdots & b_{nn} \end{vmatrix}$$

where c_{ik} is given by (10'). Developing this determinant by the first column, the resulting $(2n - 1)$ th order determinant again by its first column, etc., the result will be the determinant on the right side of (10) multiplied by $(-1)^n$ and, if n is odd, by another -1 because then the cofactor of the element -1 in the first column is equal to minus one times the minor of this element. Hence, if n is odd, that determinant is multiplied by the factor $(-1)^{n+1} = +1$, and, if n is even, by $(-1)^n = +1$. This completes the proof.

Example 8.

$$\begin{vmatrix} 10 & 14 & -2 & 1 & -1 \\ 3 & 0 & 4 \end{vmatrix} = \begin{vmatrix} 9 & 4 & 18 \\ 46 & 10 & 76 \\ 29 & 7 & 49 \end{vmatrix}$$

The determinant of the elements of an n -rowed square matrix $A = (a_{ij})$ is called the **determinant of the matrix A** and will be denoted by $\det A$; thus

$$\det A = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

Since the multiplication of determinants is performed in the same way as the multiplication of matrices, it follows that for any n -rowed square matrices A and B ,

$$(12) \quad \det(AB) = \det A \det B.$$

Problems

Evaluate the following determinants.

$$1. \begin{vmatrix} 4 & 8 \\ -1 & 2 \end{vmatrix}$$

$$2. \begin{vmatrix} 3 & 11 \\ 0 & 7 \end{vmatrix}$$

$$3. \begin{vmatrix} -2 & 3 \\ -4 & 5 \end{vmatrix}$$

$$4. \begin{vmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{vmatrix}$$

$$5. \begin{vmatrix} 3 & 1 & 0 \\ 0 & 6 & 4 \\ 0 & 0 & 1 \end{vmatrix}$$

$$6. \begin{vmatrix} 12 & 27 & 12 \\ 28 & 18 & 24 \\ 70 & 15 & 40 \end{vmatrix}$$

$$7. \begin{vmatrix} -1 & 3 & 4 \\ 3 & 4 & -1 \\ 4 & -1 & 3 \end{vmatrix}$$

$$8. \begin{vmatrix} 1 & p & p^2 \\ 1 & q & q^2 \\ 1 & r & r^2 \end{vmatrix}$$

$$9. \begin{vmatrix} 1 & 3 & 1 & -2 \\ 2 & 0 & -2 & -4 \\ -1 & 1 & 0 & 1 \\ 2 & 5 & 3 & 6 \end{vmatrix}$$

$$10. \begin{vmatrix} -4 & 1 & 4 & -2 \\ -5 & -8 & 4 & -2 \\ 8 & 16 & -6 & 15 \\ -2 & 8 & -3 & 7 \end{vmatrix}$$

$$11. \begin{vmatrix} b^2 + c^2 & a^2 & a^2 \\ b^2 & c^2 + a^2 & b^2 \\ c^2 & c^2 & a^2 + b^2 \end{vmatrix}$$

$$12. \begin{vmatrix} b^2 + c^2 & ab & ca \\ ab & c^2 + a^2 & bc \\ ca & bc & a^2 + b^2 \end{vmatrix}$$

$$13. \begin{vmatrix} a - b & m - n & x - y \\ b - c & n - p & y - z \\ c - a & p - m & z - x \end{vmatrix}$$

$$14. \begin{vmatrix} a^2 + 1 & ab & ac \\ ab & b^2 + 1 & bc \\ ac & bc & c^2 + 1 \end{vmatrix}$$

15. Write the product of the determinants in Probs. 5 and 7 as a determinant.

Verify (12) for the matrices:

16. In Prob. 11, Sec. 7.2

17. In Prob. 12, Sec. 7.2

18. Prove that the definition of a determinant D of order n yields the same value of D no matter which column is chosen [cf. (4b)]. Hint: model your work somewhat after that for rows given in the text.

19. Carry out the proof of Theorem 2 in detail.

20. Carry out the proof of Theorem 4 in detail.

7.4 Submatrices. Rank

Any matrix obtained by omitting some rows and columns from a given $(m \times n)$ matrix \mathbf{A} is called a **submatrix** of \mathbf{A} .

As a matter of convention and convenience the notion "submatrix" includes \mathbf{A} itself (as the matrix obtained from \mathbf{A} by omitting no rows or columns).

Example 1. The matrix

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

contains three (2×2) submatrices, namely,

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \begin{pmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{pmatrix}, \quad \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix},$$

two (1×3) submatrices (the two row vectors), three (2×1) submatrices (the column vectors), six (1×2) submatrices, namely,

$$(a_{11} \ a_{12}), \quad (a_{11} \ a_{13}), \quad (a_{12} \ a_{13}), \\ (a_{21} \ a_{22}), \quad (a_{21} \ a_{23}), \quad (a_{22} \ a_{23}),$$

and six (1×1) submatrices, $(a_{11}), (a_{12}), \dots, (a_{23})$.

A matrix \mathbf{A} is said to be of **rank r** if it contains at least one r -rowed square submatrix with nonvanishing determinant, while the determinant of any square submatrix having $r + 1$ or more rows, possibly contained in \mathbf{A} , is zero.

It is clear that the rank r of an $(m \times n)$ matrix can at most be equal to the smaller of the numbers m and n , but it may be less.

An $(n \times n)$ matrix \mathbf{A} has rank $r < n$ if, and only if, $\det \mathbf{A} = 0$; in this case \mathbf{A} is called singular. The matrix has rank $r = n$ if, and only if, $\det \mathbf{A} \neq 0$ and is then said to be nonsingular.

Rank $\mathbf{A} = 0$ means that \mathbf{A} is a null matrix.

The importance of the notion of rank will become obvious from our consideration of systems of linear equations in Secs. 7.6–7.8.

Example 1. The matrix

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & -1 \\ -3 & 3 & 0 \\ 2 & 2 & 4 \end{pmatrix}$$

is singular because $\det \mathbf{A} = 0$. Its rank is 2 because, for example, a two-rowed square submatrix with nonzero determinant is

$$\begin{pmatrix} 1 & -2 \\ -3 & 3 \end{pmatrix}.$$

Example 2. The matrix

$$\begin{pmatrix} 4 & 2 & 1 & 3 \\ 6 & 3 & 4 & 7 \\ 2 & 1 & 0 & 1 \end{pmatrix}$$

contains the four (3×3) submatrices

$$\begin{pmatrix} 4 & 2 & 1 \\ 6 & 3 & 4 \\ 2 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 4 & 2 & 3 \\ 6 & 3 & 7 \\ 2 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 4 & 1 & 3 \\ 6 & 4 & 7 \\ 2 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 3 \\ 3 & 4 & 7 \\ 1 & 0 & 1 \end{pmatrix}$$

whose determinants are zero. Hence $\text{rank } A < 3$. In fact, $\text{rank } A = 2$ because a (2×2) submatrix with nonzero determinant is

$$\begin{pmatrix} 4 & 1 \\ 6 & 4 \end{pmatrix}.$$

Theorem 1. *The rank of the transpose of a matrix is the same as that of the original matrix.*

Proof. Let $\text{rank } A = r$, and let R be an r -rowed square submatrix of A with $\det R \neq 0$. Clearly R^T is a submatrix of A^T . By Theorem 1 in the last section, $\det R^T = \det R$. Thus $\text{rank } A^T \geq r$. On the other hand, if A contains an $(r+1)$ -rowed square submatrix S , then, by definition of rank, $\det S = 0$. Since S corresponds to S^T in A^T , and $\det S^T = 0$, it follows that A^T cannot contain an $(r+1)$ -rowed square submatrix with nonzero determinant. Hence $\text{rank } A^T \leq r$. Altogether $\text{rank } A^T = r$, and the proof is complete.

Theorem 2. *The rank of a matrix equals the maximum number of linearly independent row vectors or column vectors.*

This will be proved in Sec. 7.8. (For linear dependence of vectors see Sec. 5.7.)

Problems

Find all the submatrices of the following matrices.

$$1. \begin{vmatrix} 2 & 0 \\ 4 & 3 \\ -1 & 5 \end{vmatrix} \quad 2. \begin{pmatrix} 4 & 8 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & -1 \end{pmatrix} \quad 3. \begin{pmatrix} 8 & 0 & 4 & 2 \\ 1 & 9 & 6 & 6 \end{pmatrix} \quad 4. \begin{vmatrix} 2 & 1 & 8 & 0 \\ 0 & 2 & 8 & 0 \\ 7 & 5 & 3 & 1 \end{vmatrix}$$

Determine the rank of the following matrices.

$$5. \begin{pmatrix} 1 & 3 \\ 3 & 9 \end{pmatrix} \quad 6. \begin{pmatrix} 3 & 1 & 2 \\ 4 & 0 & 6 \end{pmatrix} \quad 7. \begin{pmatrix} 7 & 1 & 0 \\ -14 & -2 & 0 \end{pmatrix}$$

$$8. \begin{pmatrix} 2 & 2 \\ 0 & 1 \\ 4 & 5 \end{pmatrix} \quad 9. \begin{pmatrix} 3 & -3 & 0 \\ 1 & 4 & 5 \\ 4 & 8 \end{pmatrix} \quad 10. \begin{pmatrix} 6 & 1 & 8 & 3 \\ 2 & 3 & 0 & 2 \\ 4 & -1 & -8 & -3 \end{pmatrix}$$

7.5 Systems of n Linear Equations in n Unknowns. Cramer's Rule

This and subsequent sections will be devoted to systems of linear algebraic equations. In this section we consider a system in which the number of equations equals the number of the unknowns. We shall see later that other cases may be reduced to this case which, therefore, is basic in this connection.

The system now under consideration will consist of n equations in n unknowns x_1, \dots, x_n :

$$(1) \quad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots &\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

Here the a_{ik} and b_i are given numbers, called the **coefficients** of the system.

The determinant

$$(2) \quad D = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

is called the **determinant of the system**.

If the b_i are all zero, the system is said to be **homogeneous**. If at least one of the coefficients b_1, \dots, b_n is not zero, the system is said to be **non-homogeneous**.

A set of numbers x_1, \dots, x_n which satisfy all the n equations (1) is called a **solution** of (1).

In the present section we shall derive a basic formula for the solutions of the system (1). This needs some preparation, as follows.

By developing D by the k th column we obtain

$$(3) \quad D = a_{1k}C_{1k} + a_{2k}C_{2k} + \cdots + a_{nk}C_{nk}$$

where C_{ik} is the cofactor of the element a_{ik} in D . If we replace the elements of the k th column of D by any other numbers, we obtain a new determinant, say, \tilde{D} . Clearly, its development by the k th column will be of the form (3), with a_{1k}, \dots, a_{nk} replaced by those new elements and the cofactors C_{ik} as before. In particular, if we choose as new elements the elements a_{1l}, \dots, a_{nl} of the l th column of D (where $l \neq k$), then the development of the resulting determinant \tilde{D} becomes

$$(4) \quad a_{1l}C_{1k} + a_{2l}C_{2k} + \cdots + a_{nl}C_{nk} = 0 \quad (l \neq k),$$

because D has two identical columns and is zero (cf. Theorem 6 in Sec. 7.3).

Developing D by the i th row, we have

$$(5) \quad D = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}.$$

If we replace the i th row of D by the j th row, without changing the other rows (including the j th), the resulting determinant is zero, and by developing it by the row which has been changed we obtain

$$(6) \quad a_{j1}C_{i1} + a_{j2}C_{i2} + \cdots + a_{jn}C_{in} = 0 \quad (j \neq i).$$

We now return to our system (1) and shall derive a formula for the solutions of the system when $D \neq 0$.

If we multiply the first equation (1) by C_{1k} , the second by C_{2k}, \dots , the last by C_{nk} and add the resulting equations, we first obtain

$$\begin{aligned} C_{1k}(a_{11}x_1 + \cdots + a_{1n}x_n) + \cdots + C_{nk}(a_{n1}x_1 + \cdots + a_{nn}x_n) \\ = b_1C_{1k} + \cdots + b_nC_{nk}. \end{aligned}$$

The expression on the left may be written

$$x_1(a_{11}C_{1k} + \cdots + a_{n1}C_{nk}) + \cdots + x_n(a_{1n}C_{1k} + \cdots + a_{nn}C_{nk}).$$

From (3) we conclude that the factor of x_k in this representation is equal to D , and from (4) it follows that the factor of x_l ($l \neq k$) is zero. Thus

$$x_kD = b_1C_{1k} + b_2C_{2k} + \cdots + b_nC_{nk}.$$

If $D \neq 0$, we may divide, finding

$$(7) \quad x_k = \frac{1}{D} (b_1C_{1k} + b_2C_{2k} + \cdots + b_nC_{nk}), \quad k = 1, \dots, n.$$

Hence, if (1) has a solution x_1, \dots, x_n , it must be given by (7).

Conversely, let us show that (7) actually is a solution of (1). By inserting (7) into the i th equation we obtain

$$\begin{aligned} a_{i1}x_1 + \cdots + a_{in}x_n \\ = \frac{1}{D} [a_{i1}(b_1C_{11} + \cdots + b_nC_{n1}) + \cdots + a_{in}(b_1C_{1n} + \cdots + b_nC_{nn})] \\ = \frac{1}{D} [b_1(a_{i1}C_{11} + \cdots + a_{in}C_{1n}) + \cdots + b_n(a_{i1}C_{n1} + \cdots + a_{in}C_{nn})]. \end{aligned}$$

From (5) we see that in this representation, b_i is multiplied by D , and from (6) it follows that b_j ($j \neq i$) is multiplied by zero. Hence the expression on the right reduces to

$$\frac{1}{D} b_i D = b_i,$$

which means that (7) satisfies (1).

Noting that the expression in parentheses in (7) is the development of the determinant obtained by replacing the k th column of D by the column with elements b_1, \dots, b_n , we may sum up our result as follows.

Cramer's theorem. *If the determinant D of the system (1) of n linear equations in the same number of unknowns is not zero, the system has precisely one solution x_1, \dots, x_n , given by the formulas*

$$(8) \quad x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \dots, \quad x_n = \frac{D_n}{D} \quad (\text{Cramer's rule})$$

where D_k is the determinant obtained from D by replacing in D the k th column by the column with the elements b_1, \dots, b_n .

Consequently, if the system is homogeneous ($b_1 = 0, \dots, b_n = 0$), then $D_1 = 0, \dots, D_n = 0$, and we obtain

Theorem 2. *A homogeneous system of n linear equations in n unknowns with nonvanishing determinant has only the trivial solution*

$$x_1 = 0, \quad x_2 = 0, \quad \dots, \quad x_n = 0.$$

On the other hand, we have

Theorem 3. *A nonhomogeneous system whose determinant is zero but $D_i \neq 0$ for at least one $i = 1, 2, \dots, n$, has no solution.*

In Sec. 7.9 we shall see that there are methods for solving systems of linear equations which require much less numerical work than Cramer's rule does.

Example 1. Solve

$$\begin{aligned} 2x_1 - x_2 + 2x_3 &= 2 \\ x_1 + 10x_2 - 3x_3 &= 5 \\ -x_1 + x_2 + x_3 &= -3. \end{aligned}$$

The determinant of the system is

$$D = \begin{vmatrix} 2 & -1 & 2 \\ 1 & 10 & -3 \\ -1 & 1 & 1 \end{vmatrix} = 46.$$

The determinants in the numerators in (8) are

$$D_1 = \begin{vmatrix} 2 & -1 & 2 \\ 5 & 10 & -3 \\ -3 & 1 & 1 \end{vmatrix} = 92, \quad D_2 = \begin{vmatrix} 2 & 2 & 2 \\ 1 & 5 & -3 \\ -1 & -3 & 1 \end{vmatrix} = 0, \quad D_3 = \begin{vmatrix} 2 & -1 & 2 \\ 1 & 10 & 5 \\ -1 & 1 & -3 \end{vmatrix} = -46.$$

Therefore, $x_1 = 2$, $x_2 = 0$, and $x_3 = -1$.

Example 2. In the case of the system

$$\begin{aligned} x + y &= 0 \\ x - y &= 0 \end{aligned}$$

the determinants occurring in (8) are $D = -2$, $D_1 = D_2 = 0$. Hence the only solution is $x = y = 0$, and our example illustrates Theorem 2.

Example 3. The nonhomogeneous system

$$x + y = 1$$

$$x + y = 0$$

has no solution, because $D = 0$, $D_1 = 1$, $D_2 = -1$, in agreement with Theorem 3.

Problems

Solve the following systems of linear equations.

$$\begin{array}{ll} 1. \quad 2x + y = 7 & 2. \quad 4x - 6y = 6 \\ 3x - 2y = 0 & 9x + 4y = -4 \end{array} \quad \begin{array}{ll} 3. \quad x + 2y = 9 & 4. \quad 4x - y = 8 \\ 2x - 3y = 4 & -x + 3y = 9 \end{array}$$

$$\begin{array}{ll} 5. \quad x + y - z = 0 & 6. \quad 2x + 2y + z = 5 \\ 2x - y + z = 3 & x + 3z = 5 \\ 4x + 2y - 2z = 2 & -x + 4y + 4z = 2 \end{array} \quad \begin{array}{ll} 7. \quad -x - 2y + z = -3 & \\ 2x + 3y - z = -2 & \\ x - y + 2z = 9 & \end{array}$$

$$\begin{array}{ll} 8. \quad 2x + 3y + 4z = 3 & 9. \quad 2x - y + 3z = 8 \\ 4y + 6z = 2 & -x + 2y + z = 4 \\ 8x + z = 9 & 3x + y - 4z = 0 \end{array} \quad \begin{array}{ll} 10. \quad 4x + 2y + z = -8 & \\ 3x - y + 2z = 4 & \\ -x - y - z = 2 & \end{array}$$

11. Using Kirchhoff's laws, show that the equations for the unknown currents I_1 , I_2 , I_3 in Fig. 205 are

$$I_1 - I_2 - I_3 = 0, \quad R_2 I_2 - R_3 I_3 = 0, \quad (R_1 + R_4) I_1 + R_2 I_2 = E_0$$

and determine these currents in terms of R_1 , R_2 , R_3 , R_4 , and E_0 .

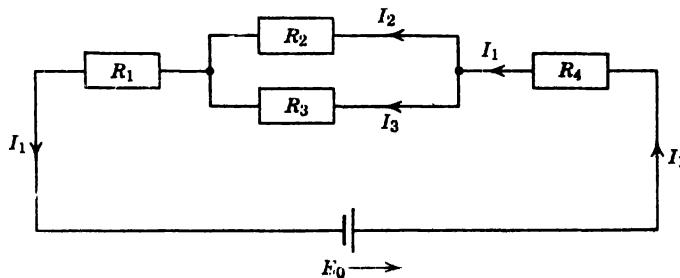
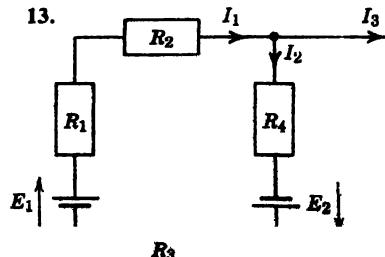
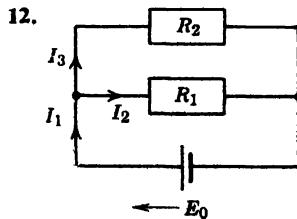


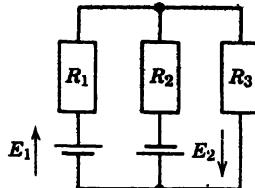
Fig. 205. Network in Problem 11.

Using Kirchhoff's laws, find the currents in the following networks.



R_5

14.



15. (Wheatstone bridge) Show that if $R_x/R_3 = R_1/R_2$ in Fig. 206, then $I = 0$. (R_0 is the resistance of the instrument by which I is measured.)

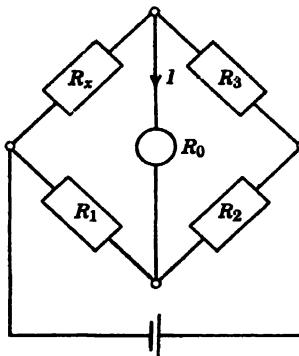


Fig. 206. Wheatstone bridge.

7.6 Arbitrary Homogeneous Systems of Linear Equations

We shall now consider homogeneous systems of linear equations where the number of equations may be different from the number of unknowns. Such a system is of the form

$$(1) \quad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \dots &\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0. \end{aligned}$$

Any set of numbers x_1, \dots, x_n that satisfies all the equations (1) is called a **solution of (1)**.

By substitution we see that the system has the **trivial solution**

$$x_1 = 0, \quad x_2 = 0, \quad \dots, \quad x_n = 0.$$

Any other solution, if it exists, is called a **nontrivial solution of (1)**.

We shall obtain conditions under which the system (1) has nontrivial solutions and develop a method for determining all solutions of the system.

Suppose that the coefficient matrix of the system

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

has rank r . We may assume $r > 0$, because $r = 0$ would mean that all the coefficients of the system are zero. Clearly, r is at most equal to the smaller of the numbers m and n . By definition of rank the matrix \mathbf{A} contains an r -rowed square matrix \mathbf{R} with nonvanishing determinant. Without specialization we may assume that the order of the equations and the order of the unknowns in each equation is chosen such that \mathbf{R} is the submatrix

$$\mathbf{R} = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \cdots & \vdots \\ \vdots & \cdots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{pmatrix}, \quad (\det \mathbf{R} \neq 0)$$

which occupies the left upper corner of the matrix \mathbf{A} . The determinant of any larger square submatrix possibly contained in \mathbf{A} is zero, by definition of rank.

Under our assumption about the order of the equations and unknowns we can now prove the following statement, which will simplify our problem.

If x_1, \dots, x_n are solutions of the first r equations of the system (1), they also satisfy the other $m - r$ equations of the system.

Proof. If $r = m$, the statement is trivial and we may assume $r < m$. Let y_1, \dots, y_n be independent variables and let

$$\begin{aligned} f_1 &= a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n \\ f_2 &= a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n \\ &\vdots \\ f_m &= a_{m1}y_1 + a_{m2}y_2 + \cdots + a_{mn}y_n. \end{aligned}$$

We now consider the determinant

$$\begin{vmatrix} a_{11} & \cdots & a_{1r} & f_1 \\ \vdots & \cdots & \vdots & \vdots \\ \vdots & \cdots & \vdots & \vdots \\ a_{r1} & \cdots & a_{rr} & f_r \\ a_{s1} & \cdots & a_{sr} & f_s \end{vmatrix} = \sum_{v=1}^n y_v \begin{vmatrix} a_{11} & \cdots & a_{1r} & a_{1v} \\ \vdots & \cdots & \vdots & \vdots \\ \vdots & \cdots & \vdots & \vdots \\ a_{r1} & \cdots & a_{rr} & a_{rv} \\ a_{s1} & \cdots & a_{sr} & a_{sv} \end{vmatrix}$$

where s is a positive integer, $r < s \leq m$, and the sum of determinants on the right is obtained by applying Theorem 4 in Sec. 7.3 to the last column of

the determinant on the left. The determinants on the right are of order $r + 1$. For $v \leq r$ they have two identical columns and are, therefore, zero. For $v > r$ they are the determinants of $(r + 1)$ -rowed square submatrices of \mathbf{A} and are zero, because $\text{rank } \mathbf{A} = r$, by assumption. Hence the determinant on the left is zero. Developing it by the last column, we have

$$f_1 D_{s1} + f_2 D_{s2} + \cdots + f_r D_{sr} + f_s \det \mathbf{R} = 0,$$

where D_{s1}, \dots, D_{sr} are determinants of certain $(r \times r)$ submatrices of \mathbf{A} (except for the sign). Since $\det \mathbf{R} \neq 0$, we may solve for f_s , finding

$$f_s = -\frac{1}{\det \mathbf{R}} (f_1 D_{s1} + \cdots + f_r D_{sr}).$$

If we write f_1, \dots, f_r, f_s at length, this becomes

$$(2) \quad a_{s1}y_1 + \cdots + a_{sn}y_n = -\frac{1}{\det \mathbf{R}} [D_{s1}(a_{11}y_1 + \cdots + a_{1n}y_n) \\ + \cdots + D_{sr}(a_{r1}y_1 + \cdots + a_{rn}y_n)].$$

Choosing for y_1, \dots, y_n a solution x_1, \dots, x_n of the first r equations of our system, we clearly have

$$a_{11}x_1 + \cdots + a_{1n}x_n = 0, \quad \dots, \quad a_{rr}x_1 + \cdots + a_{rn}x_n = 0,$$

the right side of (2) is zero, and, therefore, also

$$a_{s1}x_1 + \cdots + a_{sn}x_n = 0, \quad (s = r + 1, r + 2, \dots, m).$$

This proves our statement.

According to the statement just proved we need only consider the first r equations of our system. We may write these equations in the form

$$(3) \quad \begin{aligned} a_{11}x_1 + \cdots + a_{1r}x_r &= -(a_{1,r+1}x_{r+1} + \cdots + a_{1n}x_n) \\ a_{21}x_1 + \cdots + a_{2r}x_r &= -(a_{2,r+1}x_{r+1} + \cdots + a_{2n}x_n) \\ \dots &\dots \\ a_{rr}x_1 + \cdots + a_{rr}x_r &= -(a_{r,r+1}x_{r+1} + \cdots + a_{rn}x_n) \end{aligned}$$

(where, if $r = n$, the expressions on the right have to be replaced by zeros).

If $r = n$, the system (3) is homogeneous, and since $\det \mathbf{R} \neq 0$, it follows from Theorem 2, Sec. 7.5, that the system has only the trivial solution

$$x_1 = 0, x_2 = 0, \dots, x_n = 0.$$

If $r < n$, then we may assign arbitrary values to the unknowns x_{r+1}, \dots, x_n on the right side and then determine the unique corresponding values x_1, \dots, x_r by Cramer's rule.

We may sum up our result as follows.

Theorem 1. If the coefficient matrix of a homogeneous system of linear equations in n unknowns has rank r , then the values of $n - r$ unknowns (which

must be such that the coefficients of the other r unknowns form a matrix of rank r) remain arbitrary, and the values of the other unknowns are then uniquely determined and can be obtained by means of Cramer's rule. If $r = n$, the system has only the trivial solution $x_1 = 0, \dots, x_n = 0$.

The steps of solving our homogeneous system (1) are now as follows.

- I. Determine the rank r of the coefficient matrix A .
- II. If $r < m$, omit $m - r$ equations such that the coefficient matrix of the remaining equations still has rank r .
- III. Write these equations in the form (3) such that the square matrix of the coefficients on the left has rank r .
- IV. Solve (3) by Cramer's rule.

Theorem 1 is fundamental from the theoretical point of view. However, the resulting method for solving (1) is of limited practical interest. In fact, in Sec. 7.9 we shall see that there are other methods which require much less numerical work.

Example 1. Solve

$$\begin{aligned} 4w + 2x + y + 3z &= 0 \\ 6w + 3x + 4y + 7z &= 0 \\ 2w + x + z &= 0. \end{aligned}$$

I. The coefficient matrix A is the matrix considered in Ex. 2, Sec. 7.4. It is of rank 2, and a (2×2) submatrix R with nonzero determinant is obtained from A by omitting the second row and the first and third columns.

II. The second equation may be omitted.

III. The remaining system may be written

$$\begin{aligned} 2x + 3z &= -4w - y \\ x + z &= -2w. \end{aligned}$$

IV. In Cramer's rule,

$$D = \begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = -1, \quad D_1 = \begin{vmatrix} -4w - y & 3 \\ -2w & 1 \end{vmatrix} = 2w - y, \quad D_2 = \begin{vmatrix} 2 & -4w - y \\ 1 & -2w \end{vmatrix} = y.$$

Hence

$$x = \frac{D_1}{D} = y - 2w, \quad z = \frac{D_2}{D} = -y.$$

For example, if we choose $w = 1$ and $y = -1$, then $x = -3$ and $z = 1$, etc.

Example 2. The coefficient matrix of the system

$$\begin{aligned} x + y + z &= 0 \\ 3x + 2y + 2z &= 0 \end{aligned}$$

has rank 2. Hence one of the unknowns is arbitrary. Choosing z arbitrarily, we find $x = 0, y = -z$. Choosing y arbitrarily, we find $x = 0, z = -y$. But x cannot be chosen arbitrarily because $x = 0$; in fact, the matrix of the coefficients of the variables y and z is

$$\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$$

and has only rank 1 but not rank 2.

According to Theorem 1, nontrivial solutions of (1) exist, if the rank r of \mathbf{A} is less than the number of unknowns. This certainly happens if $m < n$, because then $r \leq m$. We thus obtain

Theorem 2. *If the number of equations of a homogeneous system of linear equations is less than the number of unknowns, the system has nontrivial solutions.*

Another immediate consequence of Theorem 1 is

Theorem 3. *A homogeneous system of n equations in the same number of unknowns has nontrivial solutions if, and only if, its coefficient determinant is zero.*

Let us use Theorem 3 to obtain some more information about the strange fact that

$$\mathbf{AB} = \mathbf{0} \quad \text{does not imply} \quad \mathbf{A} = \mathbf{0} \text{ or } \mathbf{B} = \mathbf{0}$$

which was illustrated in Sec. 7.2 by the example

$$\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

We see that both matrices on the left are singular. In fact, the following theorem holds.

Theorem 4. *If \mathbf{A} and \mathbf{B} are n -rowed square matrices, not null matrices, such that $\mathbf{AB} = \mathbf{0}$, then both \mathbf{A} and \mathbf{B} are singular.*

Proof. Let $\mathbf{b}_1, \dots, \mathbf{b}_n$ be the column vectors of \mathbf{B} . Then we may write $\mathbf{B} = (\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_n)$. From (13) in Sec. 7.2 and $\mathbf{AB} = \mathbf{0}$ we have

$$\mathbf{AB} = (\mathbf{Ab}_1 \ \dots \ \mathbf{Ab}_n) = (\mathbf{0} \ \dots \ \mathbf{0}),$$

that is

$$(4) \quad \mathbf{Ab}_k = \mathbf{0}, \quad k = 1, \dots, n$$

where $\mathbf{0}$ designates a column vector of the $(n \times n)$ null matrix. Since \mathbf{B} is not a null matrix, $\mathbf{b}_k \neq \mathbf{0}$ for at least one k . Regarding the corresponding vector equation (4) as a system of n linear equations in the "unknowns" b_{1k}, \dots, b_{nk} , the components of \mathbf{b}_k , we conclude that \mathbf{A} must be singular, because this is a necessary and sufficient condition for the existence of a nontrivial solution (Theorem 3). The proof that \mathbf{B} must be singular follows by applying the same argument to the transposed equation

$$\mathbf{B}^T \mathbf{A}^T = \mathbf{0},$$

obtained from $\mathbf{AB} = \mathbf{0}$ by the use of (12) in Sec. 7.2.

By interchanging the factors in the preceding example we obtain

$$\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Hence, if both \mathbf{A} and \mathbf{B} are singular and $\mathbf{AB} = \mathbf{0}$, it does *not* follow that $\mathbf{BA} = \mathbf{0}$.

From the proof of Theorem 4 we immediately have

Theorem 5. *If $\mathbf{AB} = \mathbf{0}$ and \mathbf{A} is nonsingular, then $\mathbf{B} = \mathbf{0}$.*

Furthermore, if $\mathbf{AB} = \mathbf{0}$ and we set $\mathbf{B} = \mathbf{S} - \mathbf{T}$, then

$$\mathbf{AB} = \mathbf{A}(\mathbf{S} - \mathbf{T}) = \mathbf{0} \quad \text{or} \quad \mathbf{AS} = \mathbf{AT}.$$

From this and Theorem 5 we obtain

Theorem 6. *The matrix equation*

$$\mathbf{AS} = \mathbf{AT} \quad \text{implies} \quad \mathbf{S} = \mathbf{T}$$

only if \mathbf{A} is nonsingular.

Problems

Find all the solutions of the following systems of linear equations.

1. $2x - y = 0$
 $x + y - z = 0$

2. $3x + 2y - 7z = 0$
 $x - y + z = 0$

3. $-4x + 3y - z = 0$
 $8x + y + 2z = 0$

4. $x + y - z = 0$
 $3x + 6y - 4z = 0$

5. $6x + 2y = 0$
 $-3x - y = 0$

6. $3x + 2y = 0$
 $7x - 4y = 0$

7. $x + y - 6z = 0$
 $-3x + y + 2z = 0$
 $x - y + 2z = 0$

8. $x + 2y - z = 0$
 $-2x - 4y + 2z = 0$
 $4x + 8y - 4z = 0$

9. $3w + 2x - y - z = 0$
 $-5x + y + 2z = 0$
 $8w + x - y - 2z = 0$

10. $-2w + 7x + y = 0$
 $3w - 3x - y - z = 0$
 $w + 5x + 2y - 5z = 0$

11. Show that the straight line through two given points $P_1: (x_1, y_1)$ and $P_2: (x_2, y_2)$ in the xy -plane is

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0.$$

12. Find the straight line through the points $(5, 2)$ and $(8, 4)$.

13. Derive from the formula in Prob. 11 the familiar formula

$$\frac{x - x_1}{x_1 - x_2} = \frac{y - y_1}{y_1 - y_2}.$$

14. Show that three points (x_1, y_1) , (x_2, y_2) , and (x_3, y_3) lie on a straight line if, and only if,

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0.$$

15. Prove Theorem 3.

7.7 Arbitrary Nonhomogeneous Systems of Linear Equations

The last remaining type of systems are nonhomogeneous systems of linear equations where the number of unknowns may be different from the number of equations. Such a system may be written in the form

$$(1) \quad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \dots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

It is clear that the solutions of this system are identical with those solutions (if they exist) of the corresponding homogeneous system

$$(2) \quad \begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n + b_1x_{n+1} &= 0 \\ a_{21}x_1 + \cdots + a_{2n}x_n + b_2x_{n+1} &= 0 \\ \dots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n + b_mx_{n+1} &= 0 \end{aligned}$$

for which $x_{n+1} = -1$. Since we know how to solve a homogeneous system, we may solve (2), set $x_{n+1} = -1$, and obtain the solutions of (1).

It may happen that (2) has only solutions for which $x_{n+1} = 0$. Then our system (1) has no solution at all.

If (2) has a solution with $x_{n+1} \neq 0$, we may multiply it by a suitable factor and obtain a solution with $x_{n+1} = -1$, which is then a solution of (1).

Necessary and sufficient conditions for the existence of a solution of (1) are as follows.

Theorem 1. *The system (1) has solutions if, and only if, the coefficient matrix A and the augmented matrix B, that is,*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \cdots & \vdots & \vdots \\ \vdots & \cdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}$$

have the same rank.

Proof. If A has rank r , it contains an $(r \times r)$ submatrix R with nonzero determinant. Since R is also a submatrix of B, the rank of B is at least r . Assume that (1) has a solution x_1^0, \dots, x_n^0 . Then

$$b_1 = a_{11}x_1^0 + a_{12}x_2^0 + \cdots + a_{1n}x_n^0, \quad \text{etc.},$$

and we may write \mathbf{B} in the form

$$\mathbf{B} = \begin{pmatrix} a_{11} & \cdots & a_{1n} & a_{11}x_1^0 + \cdots + a_{1n}x_n^0 \\ \vdots & \cdots & \vdots & \vdots \\ \vdots & \cdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & a_{m1}x_1^0 + \cdots + a_{mn}x_n^0 \end{pmatrix}.$$

Consequently, if \mathbf{B} contains $(r+1)$ -rowed square submatrices, these are either submatrices of \mathbf{A} , or their determinants can be written as sums of $(r+1)$ th order determinants of submatrices of \mathbf{A} and $(r+1)$ th order determinants with two identical columns. The latter determinants are certainly zero, and the former are zero because \mathbf{A} has rank r . Hence \mathbf{B} does not contain $(r+1)$ -rowed square submatrices with nonvanishing determinants. Therefore, its rank is less than $r+1$, that is, it is r .

Conversely, assume $\text{rank } \mathbf{A} = \text{rank } \mathbf{B} = r$. Then certainly $r \leq m$. As in Sec. 7.6, we may assume that the equations of (1) and the unknowns x_1, \dots, x_n are arranged so that the submatrix

$$(3) \quad \mathbf{R} = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \cdots & \vdots \\ \vdots & \cdots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{pmatrix}$$

has a nonvanishing determinant. Then, by Theorem 1 in Sec. 7.6, the homogeneous system (2) has a solution for which $x_{r+1}, \dots, x_n, x_{n+1}$ remain arbitrary. Hence we may choose $x_{n+1} = -1$ and obtain a solution of (1). This completes the proof.

The steps of solving our nonhomogeneous system (1) are quite similar to those in the homogeneous case (Sec. 7.6):

- I. Determine the ranks of \mathbf{A} and \mathbf{B} . If $\text{rank } \mathbf{B} \neq \text{rank } \mathbf{A}$, the system has no solution. If $\text{rank } \mathbf{B} = \text{rank } \mathbf{A} = r$, the further steps are:
- II. If $\text{rank } r < m$, omit $m-r$ equations such that the coefficient matrix of the remaining r equations still has rank r .
- III. Write the remaining equations in a form similar to (3) in Sec. 7.6, that is, assuming that the equations and variables are arranged so that the matrix \mathbf{R} , given by (3), this section, has a nonvanishing determinant, write

$$(4) \quad \begin{aligned} a_{11}x_1 + \cdots + a_{1r}x_r &= b_1 - (a_{1,r+1}x_{r+1} + \cdots + a_{1n}x_n) \\ a_{21}x_1 + \cdots + a_{2r}x_r &= b_2 - (a_{2,r+1}x_{r+1} + \cdots + a_{2n}x_n) \\ \dots &\dots \\ a_{r1}x_1 + \cdots + a_{rr}x_r &= b_r - (a_{r,r+1}x_{r+1} + \cdots + a_{rn}x_n) \end{aligned}$$

(where, if $r = n$, the expressions on the right are b_1, \dots, b_r).

- IV. Solve (4) by Cramer's rule.

As in Sec. 7.6 we should mention that, although Theorem 1 is of basic *theoretical* importance, the *practical* value of the present method for solving (1) is small; an important practical method of that type will be considered in Sec. 7.9.

Example 1. Solve

$$3x_1 + 2x_2 + 2x_3 - 5x_4 = 8$$

$$2x_1 + 5x_2 + 5x_3 - 18x_4 = 9$$

$$4x_1 - x_2 - x_3 + 8x_4 = 7$$

I. The four (3×3) submatrices in the coefficient matrix A have vanishing determinants, while the submatrix obtained by omitting from A the last row and the last two columns is

$$R = \begin{pmatrix} 3 & 2 \\ 2 & 5 \end{pmatrix}, \quad \text{and} \quad \det R = 11$$

so that $\text{rank } A = 2$. Similarly, the augmented matrix has rank 2. Hence the system has solutions.

II. Omit the last equation, since this corresponds to the last row of A , the row which does not occur in R .

III. The reduced system, written in the form (4), becomes

$$3x_1 + 2x_2 = 8 - 2x_3 + 5x_4$$

$$2x_1 + 5x_2 = 9 - 5x_3 + 18x_4$$

IV. Cramer's rule yields

$$x_1 = 2 - x_4, \quad x_2 = 1 - x_3 + 4x_4,$$

where x_3 and x_4 are arbitrary.

Problems

Solve the following systems of linear equations.

1. $3x - 7y + 35z = 18$
 $5x + 4y - 20z = -17$

2. $x - 4y + z = -2$
 $3x - 20y + 5z = -4$

3. $-3x + 3y + 7z = 4$
 $x - y + 2z = 3$

4. $-14x + 8y + 4z = 12$
 $32x + 4y + 2z = 6$

5. $6x - 3y + 15z = -15$
 $x + 3y - 8z = 15$

6. $x - 3y + 2z = 2$
 $5x - 15y + 7z = 10$

7. $7w + 2x - 2y - 6z = 6$
 $w + 3x - 3y - 9z = -10$

8. $2w - 2x - 12y + 6z = 4$
 $3w - 3x - 2y + z = -2$

9. $5w + 2x - 2y + 4z = 17$
 $-3w + x - y + 2z = -8$

10. $-5w + 25x + 5y + 7z = 45$
 $3w - 15x - 3y - 11z = -27$

11. $3x + y + z = 8$
 $-x + y - 2z = -5$
 $2x + 2y + 2z = 12$
 $-2x + 2y - 3z = -7$

12. $-2x + 2y - 3z = -2$
 $-x + y - 2z = 6$
 $-x - y - z = -1$
 $3x + y + z = -1$

13. $w - 2x + 13y - 3z = -3$
 $-3w + x - 8y + 2z = 2$
 $4w + 3x - 9y + z = 1$

14. $w - 2x + 5y - 3z = 0$
 $4w - 8x + 6y - 2z = 6$
 $-3w + 6x + y + z = 0$

7.8 Further Properties of Systems of Linear Equations

We shall now consider some further basic properties of systems of linear equations. To simplify our notation we may write such a system

$$(1^*) \quad \begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

as a single vector equation. In fact, from the definition of matrix multiplication it follows that we may write (1^{*}) in the form

$$(1) \quad \mathbf{Ax} = \mathbf{b}$$

where \mathbf{A} is the coefficient matrix of the system and \mathbf{x} and \mathbf{b} are column vectors with components x_1, \dots, x_n and b_1, \dots, b_m , respectively.

If the system is homogeneous, $\mathbf{b} = \mathbf{0}$; otherwise, $\mathbf{b} \neq \mathbf{0}$.

The solutions of the homogeneous system have the following simple but important property.

Theorem 1. *If $\mathbf{x}_{(1)}$ and $\mathbf{x}_{(2)}$ are solutions of the homogeneous system*

$$(2) \quad \mathbf{Ax} = \mathbf{0},$$

then $\mathbf{x} = c_1\mathbf{x}_{(1)} + c_2\mathbf{x}_{(2)}$, where c_1 and c_2 are any constants, is a solution of the homogeneous system.

Proof. By assumption, $\mathbf{Ax}_{(1)} = \mathbf{0}$, $\mathbf{Ax}_{(2)} = \mathbf{0}$. Hence

$$\mathbf{Ax} = \mathbf{A}(c_1\mathbf{x}_{(1)} + c_2\mathbf{x}_{(2)}) = c_1\mathbf{Ax}_{(1)} + c_2\mathbf{Ax}_{(2)} = \mathbf{0}.$$

Note that the theorem *does not hold* for nonhomogeneous systems.

p solutions $\mathbf{x}_{(1)}, \dots, \mathbf{x}_{(p)}$ of the homogeneous system (2) are called *linearly dependent* or *independent*, according to whether these p vectors are linearly dependent or independent in the sense of the usual definition (cf. Sec. 5.7).

We first prove the following useful general criterion for linear dependence and independence of vectors.

Theorem 2. *p vectors $\mathbf{x}_{(1)}, \dots, \mathbf{x}_{(p)}$ are linearly independent if the matrix with column vectors $\mathbf{x}_{(1)}, \dots, \mathbf{x}_{(p)}$ has rank p ; they are linearly dependent if that rank is less than p .*

Proof. Those vectors are linearly dependent if there are numbers k_1, \dots, k_p , not all zero, such that

$$k_1\mathbf{x}_{(1)} + \cdots + k_p\mathbf{x}_{(p)} = \mathbf{0}.$$

Denoting the components of the vector $\mathbf{x}_{(j)}$ by x_{1j}, \dots, x_{nj} , we may write this vector equation in the form

$$\begin{aligned}x_{11}k_1 + \cdots + x_{1p}k_p &= 0 \\x_{21}k_1 + \cdots + x_{2p}k_p &= 0 \\&\dots \\x_{n1}k_1 + \cdots + x_{np}k_p &= 0\end{aligned}$$

In this system we may regard k_1, \dots, k_p as unknowns. From Theorem 1 in Sec. 7.6 it follows that the system has a nontrivial solution k_1, \dots, k_p if, and only if, the coefficient matrix has rank less than p . From this the statements follow.

An immediate consequence of Theorem 2 is

Theorem 3. *p vectors with $n < p$ components are always linearly dependent.*

In fact, then the coefficient matrix of the system considered just before has rank $r \leq n < p$, and this proves Theorem 3.

By combining Theorem 2 and Theorem 1 in Sec. 7.4, we also have

Theorem 4. *The rank of a matrix equals the maximum number of linearly independent column vectors or row vectors.*

Returning to our homogeneous system (2), we shall now introduce the following notion.

*A set of linearly independent solutions $\mathbf{x}_{(1)}, \dots, \mathbf{x}_{(p)}$ of the homogeneous system (2) is called a **fundamental system** of solutions of (2), if every solution \mathbf{x} of (2) can be written as a linear combination of these vectors, that is, in the form*

$$\mathbf{x} = c_1\mathbf{x}_{(1)} + \cdots + c_p\mathbf{x}_{(p)}$$

where c_1, \dots, c_p are suitable constants.

Theorem 5. *If the coefficient matrix \mathbf{A} of the homogeneous system (2) in n unknowns has rank $r (< n)$, the system has a fundamental system of solutions, consisting of $n - r$ independent solutions.*

Proof. If \mathbf{A} has rank r , then $n - r$ unknowns, say, the unknowns x_{r+1}, \dots, x_n remain arbitrary. We choose these $n - r$ unknowns in $n - r$ different ways, namely,

$$x_{r+1} = 1, \text{ and the others zero,}$$

$$(3) \quad x_{r+2} = 1, \text{ and the others zero,}$$

$$\dots$$

$$x_n = 1, \text{ and the others zero.}$$

In each case the unknowns x_1, \dots, x_r are then uniquely determined (cf. Theorem 1 in Sec. 7.6).

In this way we obtain altogether $n - r$ solutions of (2). We denote these solutions by $\mathbf{y}_{(1)}, \dots, \mathbf{y}_{(n-r)}$. To show that they are linearly independent we form the matrix having these solutions as row vectors:

$$\begin{pmatrix} \cdots & 1 & 0 & \cdots & 0 \\ \cdots & 0 & 1 & \cdots & 0 \\ \cdots & . & . & \cdots & . \\ \cdots & 0 & 0 & \cdots & 1 \end{pmatrix};$$

here the dots in the left-hand part denote the values of the unknowns x_1, \dots, x_r in these solutions. This matrix has rank $n - r$ because the determinant of the $(n - r)$ -rowed square submatrix formed of the last $n - r$ columns of the matrix is not zero. This proves linear independence of $\mathbf{y}_{(1)}, \dots, \mathbf{y}_{(n-r)}$.

We finally show that these solutions form a fundamental system. We have to prove that any solution \mathbf{x} of (2) can be represented as a linear combination of the vectors $\mathbf{y}_{(1)}, \dots, \mathbf{y}_{(n-r)}$. For this purpose we introduce the vector

$$(4) \quad \mathbf{z} = \mathbf{x} - x_{r+1}\mathbf{y}_{(1)} - x_{r+2}\mathbf{y}_{(2)} - \cdots - x_n\mathbf{y}_{(n-r)}$$

and denote its components by z_1, \dots, z_n . From the form of the last $n - r$ components of the vectors $\mathbf{y}_{(j)}$ [cf. (3)], it follows that the last $n - r$ components of \mathbf{z} are zero:

$$(5) \quad z_{r+1} = x_{r+1} - x_{r+1} \cdot 1 = 0, \quad \dots, \quad z_n = x_n - x_n \cdot 1 = 0.$$

Since the vectors on the right side of (4) are solutions of (2), it follows from Theorem 1 that \mathbf{z} is a solution of (2). Now (2) may be written

$$(6) \quad \mathbf{a}_{(1)}x_1 + \cdots + \mathbf{a}_{(n)}x_n = \mathbf{0},$$

where $\mathbf{a}_{(j)}$ is a column vector with components a_{1j}, \dots, a_{mj} . By inserting the solution \mathbf{z} and using (5) we obtain

$$(7) \quad \mathbf{a}_{(1)}z_1 + \cdots + \mathbf{a}_{(r)}z_r = \mathbf{0}.$$

From the assumptions that \mathbf{A} has rank r and x_{r+1}, \dots, x_n are the unknowns which remain arbitrary, it follows that the matrix with column vectors $\mathbf{a}_{(1)}, \dots, \mathbf{a}_{(r)}$ has rank r . Hence, by Theorem 2 (with $p = r$), these vectors are linearly independent. From this and (7) we conclude that

$$z_1 = 0, \dots, z_r = 0.$$

From this and (5) we have $\mathbf{z} = \mathbf{0}$, and this completes the proof, because (4) now yields the desired representation of the arbitrary solution \mathbf{x} :

$$\mathbf{x} = x_{r+1}\mathbf{y}_{(1)} + \cdots + x_n\mathbf{y}_{(n-r)}.$$

We now turn to nonhomogeneous systems of linear equations and prove

Theorem 6. *If the nonhomogeneous system (1) has solutions \mathbf{x} , then all these solutions are of the form*

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_h$$

where \mathbf{x}_0 is any fixed solution of (1) and \mathbf{x}_h runs through all the solutions of the corresponding homogeneous system (2).

Proof. Let \mathbf{x} be any given solution of (1) and \mathbf{x}_0 an arbitrarily chosen solution of (1). Then $\mathbf{Ax} = \mathbf{b}$, $\mathbf{Ax}_0 = \mathbf{b}$ and, therefore,

$$\mathbf{A}(\mathbf{x} - \mathbf{x}_0) = \mathbf{Ax} - \mathbf{Ax}_0 = \mathbf{0}.$$

This shows that the difference $\mathbf{x} - \mathbf{x}_0$ between any solution \mathbf{x} of (1) and any fixed solution \mathbf{x}_0 of (1) is a solution of (2), say, \mathbf{x}_h . Hence all solutions of (1) are obtained by letting \mathbf{x}_h run through all the solutions of the homogeneous system (2), and the proof is complete.

Problems

1. Show that in Theorem 2, "column vectors" may be replaced by "row vectors."

In each case find whether the given vectors are linearly dependent or independent.

- | | |
|------------------------------------|--|
| 2. (1, 2), (−3, 5) | 3. (0, 6), (3, 1), (4, 2) |
| 4. (0, 0), (1, 2) | 5. (0, 2, 3), (−1, 0, 4), (2, 2, 7) |
| 6. (4, −1, 3), (−8, 2, −6) | 7. (2, 4, 0), (7, −1, 3) |
| 8. (0, 0, 1), (3, 2, 1) | 9. (1, 2, 4), (−2, 0, 1), (−1, 2, 5) |
| 10. (4, 1, 0, −2), (−12, −3, 0, 6) | 11. (1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0) |
| 12. (1, −2, 4, 0), (−2, 4, −8, 0) | 13. (3, 2, 7), (2, 4, 1), (1, −2, 6) |

For what values of k are the following vectors linearly dependent?

14. (1, 2, 3), (4, 5, 6), (7, 8, k)
 15. (1, 2, 3, 4), (5, 6, 7, 8), (9, 10, 11, 12), (13, 14, 15, k)

7.9 Gauss's Elimination Method

We have seen that any system of linear algebraic equations can be solved by the use of determinants; the steps of solving such a system are explained on pp. 389 and 393. The last step requires the application of Cramer's rule involving determinants.

Now, if the order of a determinant is large, its evaluation is very tedious. In fact, if an n th-order determinant is evaluated by developing it by a row or column, then developing each cofactor in the same fashion, etc., $n!$ multiplications would be required. If these multiplications could be carried out

with an electronic computer which can perform 10,000 multiplications per second, the time required would be as follows:

n	5	10	12	14	16	18	20	30
$n! \approx$	10^2	$3 \cdot 10^6$	$5 \cdot 10^8$	$9 \cdot 10^{10}$	$2 \cdot 10^{18}$	$6 \cdot 10^{15}$	$2 \cdot 10^{18}$	$3 \cdot 10^{32}$
Time	0.01 sec	5 min	10 hours	3 months	60 years	20,000 years	107 years	10^{21} years

Note for comparison that several different physical arguments (arising from nuclear physics, the theory of radioactive decomposition, etc.) lead to the conclusion that the age of the universe cannot be greater than 10^{10} years.

Since the problem of solving linear equations is of great practical importance, various mathematicians have tried to develop simpler and less time-consuming procedures, and various methods for solving systems of linear equations

$$(1) \quad \begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ \dots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

have been suggested. One of the most important methods is the so-called **Gauss algorithm**¹, which is a systematic elimination method. In this method for solving (1) we proceed stepwise as follows.

First step. Elimination of x_1 from the second, third, ..., m -th equation. We may assume that the order of the equations and the order of the unknowns in each equation are such that $a_{11} \neq 0$. The variable x_1 can then be eliminated from the second, ..., m th equation by subtracting

a_{21}/a_{11} times the first equation from the second equation,

a_{31}/a_{11} times the first equation from the third equation,

etc. This gives a new system of equations of the form

$$(2) \quad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ c_{22}x_2 + \cdots + c_{2n}x_n &= b_2^* \\ &\vdots \\ c_{m2}x_2 + \cdots + c_{mn}x_n &= b_m^*. \end{aligned}$$

Any solution of (1) is a solution of (2) and conversely, because each equation of (2) was derived from two equations of (1) and, by reversing the process, (1) may be obtained from (2).

¹ We present the Gauss algorithm in its classical form; further reduction of the amount of computational work was obtained by Banachiewicz and Cholesky. These and other methods are discussed in Ref. [D1].

Second step. Elimination of x_3 from the third, \dots , m -th equation in (2). If the coefficients c_{22}, \dots, c_{mn} in (2) are not all zero, we may assume that the order of the equations and the unknowns is such that $c_{22} \neq 0$. Then we may eliminate x_3 from the third, fourth, \dots , m th equation of (2) by subtracting

c_{32}/c_{22} times the second equation from the third equation,

c_{42}/c_{22} times the second equation from the fourth equation, etc.

The further steps are now obvious. In the third step we eliminate x_3 , in the fourth step we eliminate x_4 , etc.

This process will terminate only when no equations are left or when the coefficients of all the unknowns in the remaining equations are zero. We then have a system of the form

$$(3) \quad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ k_{r1}x_1 + \cdots + k_{rn}x_n &= \tilde{b}_r \end{aligned}$$

where either $r = m$ or $r < m$. If $r < m$, the remaining equations have the form

$$0 = \tilde{b}_{r+1}, \dots, 0 = \tilde{b}_m,$$

and the system has no solution, unless $\tilde{b}_{r+1} = 0, \dots, \tilde{b}_m = 0$. If the system has a solution, we may obtain it by choosing values at pleasure for the unknowns x_{r+1}, \dots, x_n , solving the last equation in (3) for x_r , the next to the last for x_{r-1} , and so on up to the line.

In the important special case when $m = n = r$, the system (3) has triangular form and there is one, and only one, solution.

Example 1. Solve

$$(4) \quad \begin{aligned} 3w + 2x + 2y - 5z &= 8 \\ 2w + 5x + 5y - 18z &= 9 \\ 4w - x - y + 8z &= 7 \end{aligned}$$

In the first step we eliminate w from the last two equations, finding

$$(5) \quad \begin{aligned} 3w + 2x + 2y - 5z &= 8 \\ \frac{1}{3}x + \frac{1}{3}y - \frac{4}{3}z &= \frac{1}{3} \\ -\frac{1}{3}x - \frac{1}{3}y + \frac{4}{3}z &= -\frac{1}{3} \end{aligned}$$

In the second step we eliminate x from the last equation in (5), finding

$$(6) \quad \begin{aligned} 3w + 2x + 2y - 5z &= 8 \\ \frac{1}{3}x + \frac{1}{3}y - \frac{4}{3}z &= \frac{1}{3} \\ 0 &= 0 \end{aligned}$$

Solving the second equation in (6) for x , we obtain

$$x = 1 - y + 4z.$$

By inserting this in the first equation in (6) and solving for w we have

$$w = 2 - z,$$

in agreement with the result at the end of Sec. 7.7 (except for notation).

Problems

Solve the following systems of linear equations by means of the Gauss algorithm.

- | | |
|---|---|
| 1. $2x + y = 4$
$5x - 2y = 1$ | 2. $3x + y = 11$
$x - y = 5$ |
| 3. $x + y = 0$
$3x - 4y = 1$ | 4. $4x + 2y = 2$
$-x + 3y = -11$ |
| 5. $4x - y + z = 0$
$x + 2y - z = 0$
$3x + y + 5z = 0$ | 6. $-x + y + 2z = 0$
$3x + 4y + z = 0$
$2x + 5y + 3z = 0$ |
| 7. $2x + y + z = 7$
$x - y + z = 0$
$4x + 2y - 3z = 4$ | 8. $4x + 6y + z = 2$
$2x + y - 4z = 3$
$3x - 2y + 5z = 8$ |
| 9. $3x - 7y + z = 4$
$x + 5y + z = 0$
$2x - y + 3z = -2$ | 10. $3x + y - 2z = -3$
$x - y + 2z = -1$
$-4x + 3y - 6z = 4$ |
| 11. $x - 2y + z = 3$
$x + 3z = 11$
$-2y + z = 1$ | 12. $-3x + y - z = 2$
$2x + 3y + z = 0$
$x + 5y + 2z = 6$ |
| 13. $3w + x + y - 8z = 2$
$-5w - 7x + y + 16z = 2$
$2w + x - y - z = -2$
$w + 3x + 2y - 13z = 4$ | 14. $w - x + 3y - 3z = 3$
$-5w + 2x - 5y + 4z = -5$
$-3w - 4x + 7y - 2z = 7$
$2w + 3x + y - 11z = 1$ |
| 15. $w + x + y = 3$
$-3w - 17x + y + 2z = 1$
$4w + 8y - 5z = 1$
$-5x - 2y + z = 1$ | 16. $2w - 2x - y + 3z = 5$
$-w - y + z = 0$
$3w - 6x + 2y + 4z = 7$
$w - 4x + 3y + z = 2$ |

✓ 7.10 The Inverse of a Matrix

Let $\mathbf{A} = (a_{ik})$ be an n -rowed square matrix and consider the linear transformation

$$(1) \quad \mathbf{y} = \mathbf{Ax}$$

where \mathbf{x} and \mathbf{y} are column vectors with n components x_1, \dots, x_n and y_1, \dots, y_n , respectively. For given $\mathbf{y} \neq 0$, the equation (1) may be regarded as a nonhomogeneous system of n linear equations in the n unknowns

x_1, \dots, x_n . The system has a unique solution if, and only if, $\det \mathbf{A} \neq 0$, and this solution is then given by Cramer's rule:

$$x_1 = D_1/D, \dots, x_n = D_n/D, \quad \text{where} \quad D = \det \mathbf{A}$$

and D_k is obtained from D by replacing the k th column of D by the column y_1, \dots, y_n . By developing D_1 by the first column, D_2 by the second column, etc., we see that the solution may be written

$$x_1 = \frac{1}{\det \mathbf{A}} (y_1 A_{11} + \dots + y_n A_{n1})$$

(2)

$$x_n = \frac{1}{\det \mathbf{A}} (y_1 A_{1n} + \dots + y_n A_{nn})$$

where A_{ik} is the cofactor of a_{ik} in $\det \mathbf{A}$. From the definition of matrix multiplication it follows that we may write (2) as a single vector equation,

$$(3) \quad \mathbf{x} = \mathbf{By}$$

where

$$(4) \quad \mathbf{B} = \frac{1}{\det \mathbf{A}} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{1n} \\ \vdots & \ddots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

note that in \mathbf{B} the cofactor A_{ik} occupies the same place as a_{ki} (not a_{ik}) does in \mathbf{A} .

Now, by inserting (3) into (1), we obtain

$$\mathbf{y} = \mathbf{A}(\mathbf{By}) = (\mathbf{AB})\mathbf{y},$$

the so-called *identity transformation*, which transforms \mathbf{y} into itself. Hence we must have

$$(5) \quad \mathbf{AB} = \mathbf{I},$$

where \mathbf{I} is the n -rowed unit matrix. Similarly, by inserting (1) into (3) we obtain

$$\mathbf{x} = \mathbf{B}(\mathbf{Ax}) = (\mathbf{BA})\mathbf{x},$$

and $\mathbf{BA} = \mathbf{I}$, as before. Altogether,

$$(6^*) \quad \mathbf{AB} = \mathbf{BA} = \mathbf{I}.$$

The matrix \mathbf{B} that satisfies (6*) is called the *inverse or reciprocal of \mathbf{A}* and will be denoted by \mathbf{A}^{-1} . Thus

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}.$$

By applying (12) in Sec. 7.3 to this equation we have

$$\det(AA^{-1}) = \det A \det A^{-1} = 1$$

which holds if, and only if, $\det A \neq 0$ and $\det A^{-1} \neq 0$. Since the solution of the system considered before is unique, A^{-1} is unique.

We may sum up our result as follows.

Theorem 1. *The inverse A^{-1} of a square matrix A exists if, and only if, A is nonsingular (that is, $\det A \neq 0$). Then the inverse A^{-1} of A is unique; it is the matrix B in (4) which is obtained from A by replacing each element a_{ik} by its cofactor A_{ik} and then transposing the resulting matrix and multiplying it by $1/\det A$.*

For a two-rowed square matrix we obtain from (4)

$$(7) \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad A^{-1} = \frac{1}{\det A} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

In the case of a diagonal matrix we simply have

$$(8) \quad A = \begin{pmatrix} 1 & \cdots & 0 \\ & \ddots & \vdots \\ & \cdots & \ddots \\ 0 & \cdots & a_{nn} \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} 1/a_{11} \\ & \ddots \\ & & 1/a_{nn} \end{pmatrix}$$

the elements of A^{-1} in the principal diagonal are the reciprocals of those of A .

Example 1.

$$A = \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}, \quad A^{-1} = \frac{1}{10} \begin{pmatrix} 4 & -1 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 0.4 & -0.1 \\ -0.2 & 0.3 \end{pmatrix}.$$

Example 2.

$$A = \begin{pmatrix} -3 & 6 & -11 \\ 3 & -4 & 6 \\ 4 & -8 & 13 \end{pmatrix}, \quad \det A = 10, \quad A^{-1} = \begin{pmatrix} -0.4 & 1.0 & -0.8 \\ -1.5 & 0.5 & -1.5 \\ -0.8 & 0 & -0.6 \end{pmatrix}.$$

From (6), with A replaced by C , we have

$$CC^{-1} = I.$$

If we now take for C the inverse A^{-1} of a given matrix A , this becomes

$$A^{-1}(A^{-1})^{-1} = I.$$

By premultiplying both sides by A we obtain

$$(9) \quad (A^{-1})^{-1} = A;$$

that is, *the inverse of the inverse is the given matrix A* .

The inverse of a product \mathbf{AC} can be obtained by inverting each factor and multiplying the results *in reverse order*:

$$(10) \quad (\mathbf{AC})^{-1} = \mathbf{C}^{-1}\mathbf{A}^{-1}.$$

To prove (10), we start from (6), with \mathbf{A} replaced by \mathbf{AC} , that is,

$$\mathbf{AC}(\mathbf{AC})^{-1} = \mathbf{I}.$$

By premultiplying this by \mathbf{A}^{-1} and using $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ we obtain

$$\mathbf{C}(\mathbf{AC})^{-1} = \mathbf{A}^{-1}.$$

If we premultiply this by \mathbf{C}^{-1} , the result follows.

Of course, (10) may be generalized to products of more than two matrices; by induction we obtain

$$(11) \quad (\mathbf{AC} \cdots \mathbf{PQ})^{-1} = \mathbf{Q}^{-1}\mathbf{P}^{-1} \cdots \mathbf{C}^{-1}\mathbf{A}^{-1}.$$

Problems

1. Check the result in Exs. 1 and 2 by verifying that $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.
2. Show that the inverse of $\mathbf{B} = k\mathbf{A}$ is $\mathbf{B}^{-1} = \mathbf{A}^{-1}/k$.

In each case find the inverse of the given matrix and check the result.

3. $\begin{pmatrix} -3 & 5 \\ 2 & 1 \end{pmatrix}$	4. $\begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}$	5. $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$
$\begin{pmatrix} 4 & -2 & 2 \\ -2 & -4 & 4 \\ -4 & 2 & 8 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
9. $\begin{pmatrix} 3 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{pmatrix}$	10. $\begin{pmatrix} 5 & -1 & 5 \\ 0 & 2 & 0 \\ -5 & 3 & -15 \end{pmatrix}$	11. $\begin{pmatrix} 4 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & -3 \end{pmatrix}$

In each case find the inverse of the given linear transformation.

12. $x^* = 19x + 2y - 9z$ $y^* = -4x - y + 2z$ $z^* = -2x + z$	13. $x^* = -2x - 2y + 7z$ $y^* = 4x + 3y - 12z$ $z^* = -x + 2z$
--	---

14. Show that the inverse of \mathbf{A}^2 is $(\mathbf{A}^{-1})^2$.
15. Find the inverse of the square of the matrices in Probs. 6 and 9.

Consider a four-terminal network (Fig. 207) in which an input signal is applied to one pair of terminals and the output signal is taken from the other pair. Assume that the network is linear, that is, the currents i_1 and i_2 are linear functions of u_1 and u_2 , say,

$$(12) \quad \begin{aligned} i_1 &= a_{11}u_1 + a_{12}u_2 \\ i_2 &= a_{21}u_1 + a_{22}u_2 \end{aligned}$$

$u_2 |$

Fig. 207. Four-terminal network.

16. Show that (12) may be written $\mathbf{i} = \mathbf{A}\mathbf{u}$, where

$$\mathbf{i} = \begin{pmatrix} i_1 \\ i_2 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

and $\mathbf{u} = \mathbf{A}^{-1}\mathbf{i}$.

17. Show that the input potential and current may be expressed as linear functions of the output potential and current, say,

$$(13) \quad \begin{aligned} u_1 &= t_{11}u_2 + t_{12}i_2, \\ i_1 &= t_{21}u_2 + t_{22}i_2, \end{aligned} \quad \text{or} \quad \mathbf{v}_1 = \mathbf{T}\mathbf{v}_2 \quad \text{where} \quad \mathbf{v}_1 = \begin{pmatrix} u_1 \\ i_1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} u_2 \\ i_2 \end{pmatrix},$$

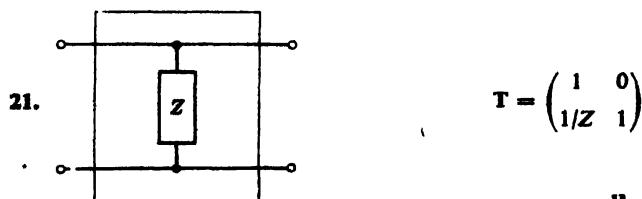
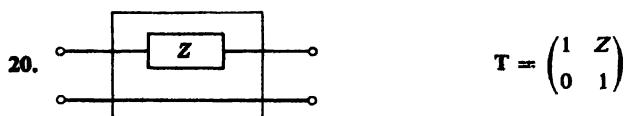
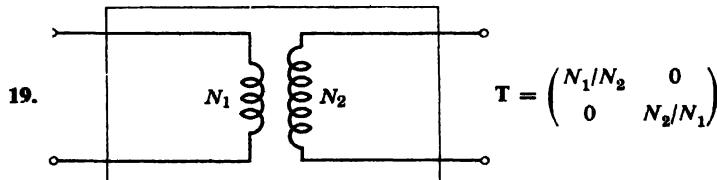
and the "transmission matrix" \mathbf{T} is

$$\mathbf{T} = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} = \frac{1}{a_{21}} \begin{pmatrix} -a_{22} & 1 \\ -\det \mathbf{A} & a_{11} \end{pmatrix}.$$

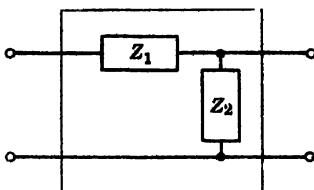
Hint: to verify the last expression, start from (12); express u_1 in terms of u_2 and i_2 ; then express i_1 in terms of u_2 and i_2 ; finally, compare the resulting representations with (13).

18. Express \mathbf{A} in terms of the elements of \mathbf{T} .

In each case show that the given transmission matrix \mathbf{T} corresponds to the indicated four-terminal network.

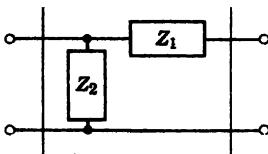


22.



$$\begin{pmatrix} 1 + Z_1/Z_2 & Z_1 \\ 1/Z_2 & 1 \end{pmatrix}$$

23.



$$T = \begin{pmatrix} 1 & Z_1 \\ 1/Z_2 & 1 + Z_1/Z_2 \end{pmatrix}$$

24. If two four-terminal networks are connected in cascade (Fig. 208), show that the resulting network may be regarded as a four-terminal network for which $v_1 = Tv_2$ where

$$v_1 = \begin{pmatrix} u_1 \\ i_1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} u_2 \\ i_2 \end{pmatrix}, \quad T = T_1 T_2.$$

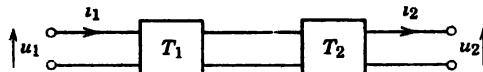


Fig. 208. Problem 24.

25. Applying the result of Prob. 24, obtain the matrix in Prob. 22 from those in Prob. 20 (with $Z = Z_1$) and Prob. 21 (with $Z = Z_2$). Obtain the matrix in Prob. 23 from those in Probs. 20 and 21.

7.11 Eigenvalues. Eigenvectors

Let $A = (a_{jk})$ be a given square n -rowed matrix and consider the vector equation

$$(1) \quad Ax = \lambda x$$

where λ is a number.

It is clear that the zero vector $x = 0$ is a solution of (1) for any value of λ . A value of λ for which (1) has a solution $x \neq 0$ is called an **eigenvalue**² or **characteristic value** of the matrix A . The corresponding solutions $x \neq 0$ of (1) are called **eigenvectors** or **characteristic vectors** of A corresponding to that eigenvalue λ . The set of the eigenvalues is called the **spectrum** of A .

The problem of determining the eigenvalues and eigenvectors of a matrix is called an **eigenvalue problem**.³ Problems of this type occur in connection

² German: Eigenwert.

³ More precisely: an algebraic eigenvalue problem, because there are other eigenvalue problems involving a differential equation (see Secs. 8.12 and 9.3) or an integral equation.

with physical and technical applications. Therefore, the student should know the fundamental ideas and concepts which are important in this field of mathematics. During the last two decades various new methods for the approximate determination of eigenvalues have been developed and other methods which have been known for still a longer time have been put into a form which is suitable for electronic computers. Cf. Ref. [D1] in Appendix 1.

Let us mention two simple problems which lead to an equation of the form (1).

Example 1. The vertical motion of the mechanical system in Fig. 209 (no damping, masses of springs neglected) is governed by the simultaneous differential equations

$$(2) \quad \begin{aligned} y_1 &= -3y_1 + 2(y_2 - y_1) \\ y_2 &= -2(y_2 - y_1) \end{aligned}$$

where $y_1(t)$ and $y_2(t)$ are the displacements of the two masses such that $y_1 = 0$, $y_2 = 0$ correspond to the positions of static equilibrium. The derivation of these equations is similar to that in Sec. 2.6. We may write the system in the form

$$\begin{aligned} y_1 &= -5y_1 + 2y_2 \\ y_2 &= 2y_1 - 2y_2 \end{aligned}$$

or as a single vector equation

$$\ddot{\mathbf{y}} = \mathbf{A}\mathbf{y},$$

where $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ and $\mathbf{A} = \begin{pmatrix} -5 & 2 \\ 2 & -2 \end{pmatrix}$.

To solve this equation we substitute

$$(3) \quad \mathbf{y} = \mathbf{x}e^{\omega t}.$$

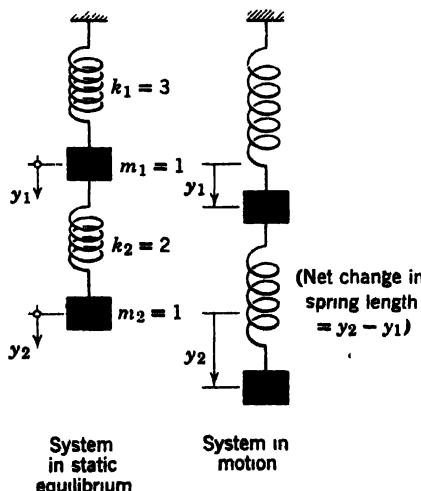


Fig. 209, Example 1.

This yields

$$\omega^2 \mathbf{x} e^{\omega t} = \mathbf{A} \mathbf{x} e^{\omega t}$$

or

$$\mathbf{A} \mathbf{x} = \lambda \mathbf{x}$$

where $\lambda = \omega^2$.

Hence, for (3) to be a solution of (2), not identically zero, $\omega^2 = \lambda$ must be an eigenvalue of \mathbf{A} and the vector \mathbf{x} in (3) must be a corresponding eigenvector.

Example 2. Given a linear transformation

$$\mathbf{y} = \mathbf{Ax} \quad (\mathbf{A} \text{ real}),$$

does there exist a real vector $\mathbf{x} \neq \mathbf{0}$ for which the corresponding vector \mathbf{y} is real and has the same direction and sense as \mathbf{x} ? This means we want to find a vector \mathbf{x} such that

$$\mathbf{y} = \mathbf{Ax} = \lambda \mathbf{x} \quad (\lambda \text{ real and positive}).$$

Clearly, such a vector exists if, and only if, \mathbf{A} has a real positive eigenvalue. \mathbf{x} is then an eigenvector corresponding to that eigenvalue.

Let us consider (1). If \mathbf{x} is any vector, then the vectors \mathbf{x} and \mathbf{Ax} will, in general, be linearly independent. If \mathbf{x} is an eigenvector, then \mathbf{x} and \mathbf{Ax} are linearly dependent; corresponding components of \mathbf{x} and \mathbf{Ax} are then proportional, the factor of proportionality being the eigenvalue λ .

We shall now demonstrate that *any n-rowed square matrix has at least 1 and at most n distinct (real or complex) eigenvalues*.

For this purpose we write (1) at length:

$$a_{11}x_1 + \cdots + a_{1n}x_n = \lambda x_1$$

$$a_{21}x_1 + \cdots + a_{2n}x_n = \lambda x_2$$

.....

$$a_{n1}x_1 + \cdots + a_{nn}x_n = \lambda x_n.$$

By transferring the terms on the right-hand side to the left-hand side we obtain

$$(4) \quad \begin{aligned} (a_{11} - \lambda)x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + \cdots + a_{2n}x_n &= 0 \\ \dots & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + (a_{nn} - \lambda)x_n &= 0. \end{aligned}$$

According to Theorem 3 in Sec. 7.6 this homogeneous system of linear equations has a nontrivial solution if, and only if, the corresponding determinant of the coefficients is zero:

$$(5) \quad D(\lambda) = \det (\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0.$$

$D(\lambda)$ is called the characteristic determinant, and (5) is called the characteristic equation corresponding to the matrix A . By developing $D(\lambda)$ we obtain a polynomial of n th degree in λ . This is called the characteristic polynomial corresponding to A .

We have thus obtained the following important result.

Theorem 1. The eigenvalues of a square matrix A are the roots of the corresponding characteristic equation (5).

An eigenvalue which is a root of m th order of the characteristic polynomial is called an eigenvalue of m -th order of the corresponding matrix.

Once the eigenvalues have been determined, the corresponding eigenvectors can be determined from the system (4). Since the system is homogeneous, it is clear that *if x is an eigenvector of A , then kx , where k is any constant, not zero, is also an eigenvector of A corresponding to the same eigenvalue*.

Example 3. Determine the eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} 5 & 4 \\ 1 & 2 \end{pmatrix}.$$

The characteristic equation

$$D(\lambda) = \begin{vmatrix} 5 - \lambda & 4 \\ 1 & 2 - \lambda \end{vmatrix} = \lambda^2 - 7\lambda + 6 = 0$$

has the roots $\lambda_1 = 6$ and $\lambda_2 = 1$. For $\lambda = \lambda_1$ the system (4) assumes the form

$$-x_1 + 4x_2 = 0$$

$$x_1 - 4x_2 = 0.$$

Thus $x_1 = 4x_2$, and

$$x_1 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$

is an eigenvector of A corresponding to the eigenvalue λ_1 . In the same way we find that an eigenvector of A corresponding to λ_2 is

$$x_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

x_1 and x_2 are linearly independent vectors.

Example 4. The matrix

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad (a, b \text{ real}, b \neq 0)$$

has the conjugate complex eigenvalues $\lambda_1 = a + ib$ and $\lambda_2 = a - ib$. Corresponding eigenvectors are

$$x_1 = \begin{pmatrix} 1 \\ i \end{pmatrix} \quad \text{and} \quad x_2 = \begin{pmatrix} 1 \\ -i \end{pmatrix}.$$

If $b = 0$, then $\lambda_1 = \lambda_2 = a$; the matrix A has just one eigenvalue (which is then of the second order), and every vector $x \neq 0$ with two components is an eigenvector of A .

Example 5. The matrix

$$A = \begin{pmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{pmatrix}$$

has the eigenvalues $\lambda_1 = 5$ and $\lambda_2 = \lambda_3 = -3$. The vector

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

is an eigenvector of A corresponding to the eigenvalue 5, and the vectors

$$\mathbf{x}_2 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{x}_3 = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$$

are two linearly independent eigenvectors of A corresponding to the eigenvalue -3 . This agrees with the fact that, for $\lambda = -3$, the matrix $A - \lambda I$ has rank 1 and so, by Theorem 5 in Sec. 7.8, a fundamental system of solutions of the corresponding system (4), viz.,

$$\begin{aligned} x_1 + 2x_2 - 3x_3 &= 0 \\ 2x_1 + 4x_2 - 6x_3 &= 0 \\ -x_1 - 2x_2 + 3x_3 &= 0 \end{aligned}$$

consists of two linearly independent vectors.

Example 5 shows that *to an eigenvalue there may correspond several linearly independent eigenvectors*. The maximum number of linearly independent eigenvectors corresponding to an eigenvalue λ is called the **multiplicity** of λ .

Problems

1. Do there exist matrices which have no eigenvalues at all?

Find the eigenvalues and eigenvectors of the following matrices.

2. $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ 3. $\begin{pmatrix} 8 & -4 \\ 2 & 2 \end{pmatrix}$ 4. $\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$ 5. $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

6. $\begin{pmatrix} -4 & 2 \\ -1 & -1 \end{pmatrix}$ 7. $\begin{pmatrix} 6 & 8 \\ 8 & -6 \end{pmatrix}$ 8. $\begin{pmatrix} -1 & 0 \\ -2 & 1 \end{pmatrix}$ 9. $\begin{pmatrix} 1.5 & 2 \\ 2 & -1.5 \end{pmatrix}$

10. $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ 11. $\begin{pmatrix} 3 & 1 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix}$ 12. $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

13. $\begin{pmatrix} -15 & 4 & 3 \\ 10 & -12 & 6 \\ 20 & -4 & 2 \end{pmatrix}$ 14. $\begin{pmatrix} -13 & 3 & -5 \\ 0 & -4 & 0 \\ 15 & -9 & 7 \end{pmatrix}$ 15. $\begin{pmatrix} 2.6 & -0.2 & 0.2 \\ 0.2 & 2.1 & 0.4 \\ 0.4 & 0.2 & 2.8 \end{pmatrix}$

16. Show that the eigenvalues of a triangular matrix A are equal to the elements of the principal diagonal of A .
17. Show that the transpose A^T has the same eigenvalues as A .
18. Verify the statement in Prob. 17 for the matrix in Prob. 13.
19. Show that if A has real elements, then the eigenvalues of A are real or complex conjugates in pairs.
20. Show that if A is a real n -rowed square matrix and n is odd, then A has at least one real eigenvalue.
21. Developing (5), show that the constant term of $D(\lambda)$ equals $\det A$.
22. Show that $\lambda = 0$ is an eigenvalue of a square matrix A if and only if A is singular.
23. Show that if A has the eigenvalues $\lambda_1, \dots, \lambda_n$, then kA has the eigenvalues $k\lambda_1, \dots, k\lambda_n$.
24. Using Prob. 23, obtain the eigenvalues in Prob. 3 from those in Prob. 6.
25. Show that if A has the eigenvalues $\lambda_1, \dots, \lambda_n$, then A^m (m a non-negative integer) has the eigenvalues $\lambda_1^m, \dots, \lambda_n^m$.
26. Show that if A has the eigenvalues $\lambda_1, \dots, \lambda_n$, then

$$k_m A^m + k_{m-1} A^{m-1} + \dots + k_1 A + k_0 I$$

has the eigenvalues

$$k_m \lambda_j^m + k_{m-1} \lambda_j^{m-1} + \dots + k_1 \lambda_j + k_0 \quad (j = 1, \dots, n).$$

27. Show that the vertical motion of the mechanical system in Fig. 210 (no damping, masses of springs neglected) is governed by the simultaneous differential equations

$$m_1 \ddot{y}_1 = -k_1 y_1 + k_2 (y_2 - y_1)$$

$$m_2 \ddot{y}_2 = -k_2 (y_2 - y_1)$$

where dots denote derivatives with respect to the time t .

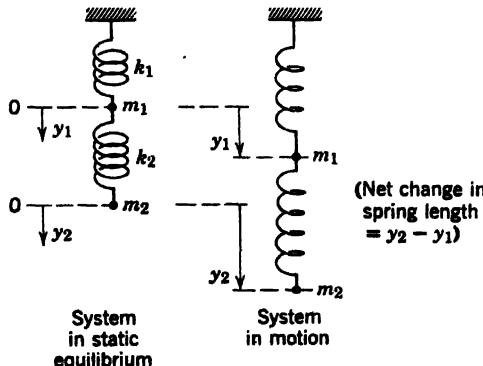


Fig. 210. Mechanical system in Problem 27.

28. Show that the system of equations in Prob. 27 may be written as a single vector equation $\ddot{\mathbf{y}} = \mathbf{Ay}$, where

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \quad \text{and} \quad \mathbf{A} = \begin{pmatrix} -\frac{k_1 + k_2}{m_1} & \frac{k_2}{m_1} \\ \frac{k_2}{m_2} & \frac{k_2}{m_2} \end{pmatrix}$$

29. To solve the equation in Prob. 28, substitute

$$y = xe^{\omega t}$$

and show that this leads to the eigenvalue problem

$$Ax = \lambda x \quad \text{where} \quad \lambda = \omega^2.$$

30. In Prob. 29, let $m_1 = m_2 = 1$, $k_1 = 3$, $k_2 = 2$. Find the eigenvalues and show that corresponding eigenvectors are

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

Conclude that this leads to the following solution of the equation in Prob. 28 (with m_1, m_2, k_1, k_2 as indicated):

$$y(t) = a_1 \mathbf{x}_1 \cos t + b_1 \mathbf{x}_1 \sin t + a_2 \mathbf{x}_2 \cos \sqrt{6}t + b_2 \mathbf{x}_2 \sin \sqrt{6}t,$$

where a_1, b_1, a_2, b_2 are arbitrary constants.

31. Show that by imposing the initial conditions $y_1(0) = 1$, $y_2(0) = 2$, $\dot{y}_1(0) = -2\sqrt{6}$, $\dot{y}_2(0) = \sqrt{6}$ or, in vector form,

$$\mathbf{y}(0) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \dot{\mathbf{y}}(0) = \begin{pmatrix} -2\sqrt{6} \\ \sqrt{6} \end{pmatrix},$$

we obtain from the solution in Prob. 30 the particular solution

$$y(t) = \mathbf{x}_1 \cos t + \mathbf{x}_2 \sin \sqrt{6}t,$$

that is,

$$y_1 = \cos t - 2 \sin \sqrt{6}t, \quad y_2 = 2 \cos t + \sin \sqrt{6}t.$$

Verify by substitution that these functions are solutions of the differential equations in Prob. 27 with $m_1 = m_2 = 1$, $k_1 = 3$, $k_2 = 2$.

32. In Prob. 30, find and graph the solutions satisfying the initial conditions $y_1(0) = 1$, $y_2(0) = 2$, $\dot{y}_1(0) = 0$, $\dot{y}_2(0) = 0$.
 33. Show that, for any positive k_1, k_2, m_1, m_2 , the eigenvalues of the matrix in Prob. 28 are negative real.

Find the solutions of the differential equations in Prob. 27, assuming that

34. $m_1 = m_2 = 1$, $k_1 = 9$, $k_2 = 6$, $y_1(0) = -2$, $y_2(0) = 1$, $\dot{y}_1(0) = 0$, $\dot{y}_2(0) = 0$.

35. $m_1 = m_2 = 1$, $k_1 = 6$, $k_2 = 4$, $y_1(0) = 0$, $y_2(0) = 0$, $\dot{y}_1(0) = \sqrt{2}$, $\dot{y}_2(0) = 2\sqrt{2}$.

In each case find the solution satisfying the given initial conditions.

36. $\dot{y}_1 = y_1 + 2y_2 \quad y_1(0) = 1, y_2(0) = 1$
 $\dot{y}_2 = -8y_1 + 11y_2$

37. $\dot{y}_1 = 3y_1 + 4y_2 \quad y_1(0) = 1, y_2(0) = 3$
 $\dot{y}_2 = 4y_1 - 3y_2$

38. $\dot{y}_1 = 4y_1 - 2y_2 \quad y_1(0) = 3, y_2(0) = 2$
 $\dot{y}_2 = y_1 + y_2$

39. $\dot{y}_1 = y_2 \quad y_1(0) = 2, y_2(0) = 0, y_3(0) = 2$
 $\dot{y}_2 = y_1 + 3y_3$
 $\dot{y}_3 = y_2$

40. $\dot{y}_1 = 3y_1 \quad y_1(0) = 2, y_2(0) = -10, y_3(0) = -27$
 $\dot{y}_2 = 5y_1 + 4y_2$
 $\dot{y}_3 = 3y_1 + 6y_2 + y_3$

7.12 Bilinear, Quadratic, Hermitian, and Skew-Hermitian Forms

An expression of the form

$$(1) \quad B = \sum_{j=1}^n \sum_{k=1}^n a_{jk} x_j y_k$$

(where the a_{jk} are numbers) is called a **bilinear form** in the $2n$ variables x_1, \dots, x_n and y_1, \dots, y_n . Writing B at length, we have

$$(1') \quad \begin{aligned} B = & a_{11}x_1y_1 + a_{12}x_1y_2 + \cdots + a_{1n}x_1y_n \\ & + a_{21}x_2y_1 + a_{22}x_2y_2 + \cdots + a_{2n}x_2y_n \\ & \dots \\ & + a_{n1}x_ny_1 + a_{n2}x_ny_2 + \cdots + a_{nn}x_ny_n. \end{aligned}$$

The n -rowed square matrix $A = (a_{jk})$ is called the *coefficient matrix* of the form. Introducing the vectors

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

we may write

$$(2) \quad B = \mathbf{x}^T A \mathbf{y}$$

where $\mathbf{x}^T = (x_1 \cdots x_n)$ is the transpose of \mathbf{x} . This follows immediately from the definition of matrix multiplication.

Example 1. If A is the unit matrix I , then

$$B = \mathbf{x}^T I \mathbf{y} = \mathbf{x}^T \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n;$$

that is, in this case, B is the scalar product of the vectors \mathbf{x} and \mathbf{y} .

If $\mathbf{y} = \mathbf{x}$, then (1) is called a **quadratic form** in the n variables x_1, \dots, x_n . Denoting this form by Q , we have

$$(3) \quad Q = \sum_{j=1}^n \sum_{k=1}^n a_{jk} x_j x_k.$$

Writing Q at length and taking corresponding terms $a_{jk}x_j x_k$ and $a_{kj}x_k x_j$ together, we obtain

$$\begin{aligned} Q = & a_{11}x_1^2 + (a_{12} + a_{21})x_1x_2 + \cdots + (a_{1n} + a_{n1})x_1x_n \\ & + a_{22}x_2^2 + \cdots + (a_{2n} + a_{n2})x_2x_n \\ & + \dots \\ & + a_{nn}x_n^2. \end{aligned}$$

We may now set

$$\frac{1}{2}(a_{jk} + a_{kj}) = c_{jk}.$$

Then $c_{kj} = c_{jk}$ and $c_{jk} + c_{kj} = a_{jk} + a_{kj}$, so that we may write

$$Q = \sum_{j=1}^n \sum_{k=1}^n c_{jk} x_j x_k.$$

If \mathbf{A} is real, then the coefficient matrix $\mathbf{C} = (c_{jk})$ in this new representation is a real symmetric matrix. This shows that any *real quadratic form* Q in n variables x_1, \dots, x_n may be written

$$Q = \mathbf{x}^T \mathbf{C} \mathbf{x}$$

where \mathbf{C} is a *real symmetric matrix*. The reader may prove that the correspondence between these forms Q and symmetric matrices \mathbf{C} is then one-to-one, that is, to each such form Q there corresponds exactly one symmetric matrix \mathbf{C} , and conversely.

Example 2. The quadratic form

$$Q = \mathbf{x}^T \mathbf{A} \mathbf{x} = 2x_1^2 + x_1 x_2 - 3x_2^2$$

has the coefficient matrix

$$\mathbf{A} = (a_{jk}) = \begin{pmatrix} 2 & 1 \\ 0 & -3 \end{pmatrix}.$$

Hence the corresponding symmetric coefficient matrix is

$$\mathbf{C} = \frac{1}{2}[\mathbf{A} + \mathbf{A}^T] = \frac{1}{2}(a_{jk} + a_{kj}) = \begin{pmatrix} 2 & \frac{1}{2} \\ \frac{1}{2} & -3 \end{pmatrix}.$$

We introduce the following notation. Let $\mathbf{A} = (a_{jk})$ be any matrix. Then $\bar{\mathbf{A}}$ denotes the matrix (\bar{a}_{jk}) which is obtained from \mathbf{A} by replacing each element a_{jk} by its *complex conjugate* \bar{a}_{jk} .

A square matrix $\mathbf{A} = (a_{jk})$ for which the transpose equals the complex conjugate, that is,

$$(4) \quad \mathbf{A}^T = \bar{\mathbf{A}} \quad (\text{that is, } a_{kj} = \bar{a}_{jk})$$

is called a **Hermitian⁴ matrix**.

From (4) we see that the elements in the principal diagonal of a Hermitian matrix are always real. Furthermore, if all the elements of a Hermitian matrix \mathbf{A} are real, then (4) assumes the form $\mathbf{A}^T = \mathbf{A}$ which means that a *real Hermitian matrix is a symmetric matrix, and so Hermitian matrices are a natural generalization of real symmetric matrices*.

A form

$$H = \bar{\mathbf{x}}^T \mathbf{A} \mathbf{x} \quad (\mathbf{A} \text{ Hermitian})$$

is called a **Hermitian form**; here the n components of the vector \mathbf{x} may be real or complex variables. Obviously, this is a generalization of a real quadratic form.

⁴ CHARLES HERMITE (1822–1901), French mathematician.

From the definition of matrix multiplication it follows that

$$H = \sum_{j=1}^n \sum_{k=1}^n a_{jk} \bar{x}_j x_k.$$

Example 3. The matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 3+i \\ 3-i & 1 \end{pmatrix}$$

is Hermitian, and the corresponding Hermitian form is

$$\begin{aligned} H &= \bar{\mathbf{x}}^T \mathbf{A} \mathbf{x} = (\bar{x}_1 \quad \bar{x}_2) \begin{pmatrix} 2 & 3+i \\ 3-i & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= 2\bar{x}_1 x_1 + (3+i)\bar{x}_1 x_2 + (3-i)\bar{x}_2 x_1 + \bar{x}_2 x_2 \\ &= 2|x_1|^2 + |x_2|^2 + 2 \operatorname{Re} [(3+i)\bar{x}_1 x_2] \end{aligned}$$

where Re denotes the real part. The last expression shows that for every choice of \mathbf{x} the value of H is a real number. We shall now prove that any Hermitian form has this remarkable property.

Theorem 1. *For every choice of \mathbf{x} the value of a Hermitian form*

$$H = \bar{\mathbf{x}}^T \mathbf{A} \mathbf{x} \tag{A Hermitian}$$

is a real number.

Proof. Using (4), we obtain

$$H = \overline{(\bar{\mathbf{x}}^T \mathbf{A} \mathbf{x})} = \mathbf{x}^T \bar{\mathbf{A}} \bar{\mathbf{x}} = \mathbf{x}^T \mathbf{A}^T \bar{\mathbf{x}}.$$

The expression on the right is a scalar. Hence transposition does not change its value. Using (12) in Sec. 7.2, we thus have

$$\mathbf{x}^T \mathbf{A}^T \bar{\mathbf{x}} = (\mathbf{x}^T \mathbf{A}^T \bar{\mathbf{x}})^T = \bar{\mathbf{x}}^T \mathbf{A} \mathbf{x} = H.$$

Hence, $H = H$, which means that H is real. This completes the proof.

A square matrix $\mathbf{A} = (a_{jk})$ for which

$$(5) \quad \mathbf{A}^T = -\bar{\mathbf{A}} \quad (\text{that is, } a_{kj} = -\bar{a}_{jk})$$

is called a **skew-Hermitian matrix**. Obviously this is a generalization of a real skew-symmetric matrix, because if all the elements of a skew-Hermitian matrix \mathbf{A} are real, then (5) assumes the form $\mathbf{A}^T = -\mathbf{A}$; that is, \mathbf{A} is then a real skew-symmetric matrix.

A form

$$(6) \quad S = \bar{\mathbf{x}}^T \mathbf{A} \mathbf{x} \tag{A skew-Hermitian}$$

is called a **skew-Hermitian form**. The reader may prove the following theorem.

Theorem 2. *For every choice of \mathbf{x} the value of a skew-Hermitian form is a purely imaginary number or zero.*

Problems

1. Verify that the right-hand side of (2) is equal to the right-hand side of (1').

Find a real symmetric matrix C such that $Q = \mathbf{x}^T C \mathbf{x}$ where Q equals

2. $3x_1^2 - 2x_1x_2 + 4x_3^2$

3. $4x_1x_2 - 5x_2^2$

4. $2(x_1 - x_3)^2$

5. $-2x_1^2 + 6x_1x_2 + 4x_3^2$

6. $5x_1^2 - 2x_1x_2 + x_3^2$

7. $(x_1 + x_2)^2 - x_3^2$

8. $(x_1 - x_2 + x_3)^2$

9. $4x_1x_3 + 2x_2x_3 + x_3^2$

10. If A and B are n -rowed Hermitian matrices and a and b are any *real* numbers, show that $C = aA + bB$ is a Hermitian matrix.
11. Show that if the variables of a quadratic form $Q = \mathbf{x}^T C \mathbf{x}$ undergo a linear transformation, say, $\mathbf{x} = P\mathbf{y}$, then $Q = \mathbf{y}^T A \mathbf{y}$ where $A = P^T C P$.
12. Consider the quadratic form $Q = \mathbf{x}^T C \mathbf{x}$ (C symmetric) in three variables x_1, x_2, x_3 . Find $\partial Q / \partial x_1, \partial Q / \partial x_2, \partial Q / \partial x_3$ and show that $\text{grad } Q = 2C\mathbf{x}$.
13. Show that any square matrix may be written as the sum of a Hermitian matrix and a skew-Hermitian matrix.
14. A real quadratic form $Q = \mathbf{x}^T C \mathbf{x}$ and its symmetric matrix $C = (c_{jk})$ are said to be **positive definite** if $Q > 0$ for all $(x_1, \dots, x_n) \neq (0, \dots, 0)$. A necessary and sufficient condition for positive definiteness is that all the determinants

$$C_1 = c_{11}, \quad C_2 = \begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix}, \quad C_3 = \begin{vmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{vmatrix}, \quad \dots, \quad C_n = \det C$$

are positive (cf. Ref. [D2]). Show that the form in Prob. 2 is positive definite.

15. Test the forms in Probs. 3 and 4 for positive definiteness.

Find $H = \mathbf{x}^T A \mathbf{x}$ where

16. $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 1+i \\ 1-i \end{pmatrix}$

17. $A = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 1 \\ i \end{pmatrix}$

18. $A = \begin{pmatrix} 3 & 1+i \\ 1-i & 2 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 2+i \\ 1+2i \end{pmatrix}$

19. $A = \begin{pmatrix} 2 & 3i \\ -3i & 4 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} i \\ -i \end{pmatrix}$

20. $A = \begin{pmatrix} -1 & 5+i \\ 5-i & 2 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

21. $A = \begin{pmatrix} a & b \\ \bar{b} & c \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 3i \\ 4 \end{pmatrix}$

22. $A = \begin{pmatrix} 1 & -i & 2i \\ i & 1 & 0 \\ -2i & 0 & 1 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 0 \\ 1 \\ i \end{pmatrix}$

23. $A = \begin{pmatrix} 1 & i & 0 \\ -i & 0 & 3 \\ 0 & 3 & 2 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

24. Show that the elements of the principal diagonal of a skew-Hermitian matrix are purely imaginary or zero.

25. Let A and B be skew-Hermitian matrices. Under what conditions is $C = aA + bB$ a skew-Hermitian matrix? (a and b are numbers.)

26. Prove Theorem 2. *Hint:* model your work somewhat after that in the proof of Theorem 1.

Find $S = \bar{x}^T A x$ where

$$27. A = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad x = \begin{pmatrix} 1 \\ i \end{pmatrix}$$

$$28. A = \begin{pmatrix} 5i & 2+i \\ -2+i & i \end{pmatrix}, \quad x = \begin{pmatrix} 2i \\ 3 \end{pmatrix}$$

$$29. A = \begin{pmatrix} 2i & 4 \\ -4 & 0 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$30. A = \begin{pmatrix} i\alpha & \mu + i\nu \\ -\mu + i\nu & i\beta \end{pmatrix}, \quad x = \begin{pmatrix} -i \\ 4 \end{pmatrix}$$

7.13 Eigenvalues of Hermitian, Skew-Hermitian, and Unitary Matrices

A square matrix $A = (a_{jk})$ for which

$$(1) \quad A^T = \bar{A}^{-1}$$

is called a **unitary matrix**.

A real unitary matrix A is called an **orthogonal matrix**. For such a matrix, (1) assumes the form

$$(2) \quad A^T = A^{-1},$$

that is, *an orthogonal matrix is a real matrix for which the inverse equals its transpose*.

A system of vectors x_1, \dots, x_n for which

$$(3) \quad \bar{x}_j^T x_k = \delta_{jk} = \begin{cases} 0 & \text{when } j \neq k \\ 1 & \text{when } j = k \end{cases} \quad (j, k = 1, \dots, n)$$

is called a **unitary system**.

Clearly, if these vectors are real, then $\bar{x}_j^T x_k = x_j^T x_k$ is the scalar product of x_j and x_k in the elementary sense, and the condition (3) means that x_1, \dots, x_n are orthogonal unit vectors. Hence a unitary system of real vectors is a system of orthogonal (mutually perpendicular) unit vectors.

Theorem 1. *The column vectors (and also the row vectors) of a unitary matrix form a unitary system.*

Proof. Let A be a unitary matrix with column vectors a_1, \dots, a_n . Then from (1) it follows that

$$A^{-1}A = \bar{A}^T A = \begin{pmatrix} \bar{a}_1^T \\ \vdots \\ \bar{a}_n^T \end{pmatrix} (a_1 \cdots a_n) = \begin{pmatrix} \bar{a}_1^T a_1 & \bar{a}_1^T a_2 & \cdots & \bar{a}_1^T a_n \\ \bar{a}_2^T a_1 & \bar{a}_2^T a_2 & \cdots & \bar{a}_2^T a_n \\ \vdots & \vdots & \ddots & \vdots \\ \bar{a}_n^T a_1 & \bar{a}_n^T a_2 & \cdots & \bar{a}_n^T a_n \end{pmatrix}$$

so that

$$\bar{\mathbf{a}}_j^T \mathbf{a}_k = \begin{cases} 0 & \text{when } j \neq k, \\ 1 & \text{when } j = k. \end{cases}$$

This shows that the column vectors of a unitary matrix form a unitary system. From (1) it follows that the inverse \mathbf{A}^{-1} of a unitary matrix \mathbf{A} is unitary (cf. Prob. 9) and, furthermore, the column vectors of \mathbf{A}^{-1} are the conjugates of the row vectors of \mathbf{A} . This shows that the row vectors of \mathbf{A} form a unitary system. Theorem 1 is proved.

With respect to orthogonal matrices, Theorem 1 means that *the row vectors (and also the column vectors) of an orthogonal matrix $\mathbf{A} = (a_{jk})$ form a system of orthogonal unit vectors, that is*

$$(4) \quad \sum_{m=1}^n a_{jm} a_{km} = \delta_{jk}.$$

For $n = 3$ these relations are identical with (7) in Sec. 5.17. This shows that *transformations of Cartesian coordinate systems into Cartesian coordinate systems (with the same unit of scale) are linear transformations whose coefficient matrices are orthogonal matrices*.

Therefore, these matrices are of great practical importance.

We shall now consider the eigenvalues of Hermitian, skew-Hermitian, and unitary matrices (Fig. 211).

Theorem 2. (a) *The eigenvalues of a Hermitian matrix are real.*
 (b) *The eigenvalues of a skew-Hermitian matrix are purely imaginary or zero.*
 (c) *The eigenvalues of a unitary matrix have the absolute value 1.*

Proof. (a) Let \mathbf{A} be Hermitian and λ an eigenvalue of \mathbf{A} . Then by definition there is a vector $\mathbf{x} \neq 0$ such that $\mathbf{Ax} = \lambda\mathbf{x}$. From this we obtain

$$\bar{\mathbf{x}}^T \mathbf{Ax} = \bar{\mathbf{x}}^T \lambda \mathbf{x} = \lambda \bar{\mathbf{x}}^T \mathbf{x}.$$

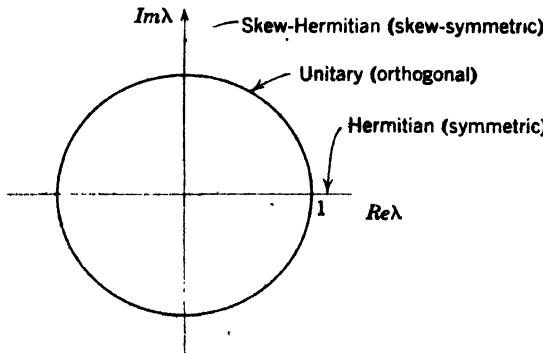


Fig. 211. Location of the eigenvalues of Hermitian, skew-Hermitian, and unitary matrices in the complex λ -plane.

Since $x \neq 0$, it follows that $\bar{x}^T x \neq 0$ and we may divide, finding

$$\lambda = \frac{\bar{x}^T A x}{\bar{x}^T x}.$$

From Theorem 1 in the last section we conclude that both the numerator and the denominator are real. Hence λ is real.

(b) Let S be skew-Hermitian, λ an eigenvalue of S , and x a corresponding eigenvector. Then $Sx = \lambda x$, and by premultiplying this by \bar{x}^T we have

$$\bar{x}^T S x = \bar{x}^T \lambda x = \lambda \bar{x}^T x.$$

Solving for λ , we obtain

$$\lambda = \frac{\bar{x}^T S x}{\bar{x}^T x}.$$

From this and Theorem 2 in the last section the proof follows.

(c) Let U be unitary, λ an eigenvalue of U , and x a corresponding eigenvector. Then

$$(5) \quad Ux = \lambda x.$$

Taking the conjugate transpose and using (12), Sec. 7.2, we obtain

$$(\bar{U}\bar{x})^T = \bar{x}^T \bar{U}^T = \bar{\lambda} \bar{x}^T.$$

Since U is unitary, $\bar{U}^T = U^{-1}$, and our equation can be written

$$(6) \quad \bar{x}^T U^{-1} = \bar{\lambda} \bar{x}^T.$$

By postmultiplying the left-hand side of (6) by the left-hand side of (5) and the right-hand side of (6) by the right-hand side of (5) we obtain

$$\bar{x}^T U^{-1} U x = \bar{\lambda} \bar{x}^T \lambda x = \lambda \bar{\lambda} \bar{x}^T x.$$

Since $U^{-1} U = I$ and $\lambda \bar{\lambda} = |\lambda|^2$, this becomes

$$\bar{x}^T x = |\lambda|^2 \bar{x}^T x.$$

Since $x \neq 0$, it follows that $\bar{x}^T x \neq 0$. Hence we must have $|\lambda|^2 = 1$. Thus $|\lambda| = 1$, and the proof is complete.

An important particular case of Theorem 2 is

Theorem 3. *The eigenvalues of a symmetric matrix are real. The eigenvalues of a skew-symmetric matrix are purely imaginary or zero. The eigenvalues of an orthogonal matrix have the absolute value 1 and are real or complex conjugates in pairs.*

Problems

Show that the following matrices are orthogonal.

$$1. \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

$$2. \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \\ 0 & 0 \end{pmatrix}$$

$$3. \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

4. Interpret the transformation $y = Ax$ geometrically, where A is the matrix in Prob. 1 and the components of x and y are Cartesian coordinates.
5. Same task as in Prob. 4, where A is the matrix in Prob. 2.
6. Show that if A is orthogonal, then $\det A = 1$ or -1 . Give examples.

Show that the following matrices are unitary and determine their eigenvalues and eigenvectors.

$$7. \begin{pmatrix} 1/\sqrt{2} & i/\sqrt{2} \\ -i/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}$$

$$8. \begin{pmatrix} 1/\sqrt{3} & i\sqrt{2}/3 \\ -i\sqrt{2}/3 & -1/\sqrt{3} \end{pmatrix}$$

9. Show that the inverse of a unitary matrix is unitary.
10. Verify the statement in Prob. 9 for the matrix in Prob. 7.
11. Show that the product of two n -rowed unitary matrices is unitary.
12. Verify the statement in Prob. 11 for the matrices in Probs. 7 and 8.
13. Show that the eigenvectors of a real symmetric matrix corresponding to different eigenvalues are orthogonal. Give an example.
14. Consider the linear transformation $y = Ax$ where A is square and set $x = T\bar{x}$, $y = T\bar{y}$ where T is nonsingular. Show that then

$$\bar{y} = \bar{A}\bar{x} \quad \text{where} \quad \bar{A} = T^{-1}AT.$$

(This transformation which transforms A into \bar{A} is called a **collineatory transformation**.)

15. Show that if $\bar{A} = T^{-1}AT$ where A is any square matrix, then $\det \bar{A} = \det A$. Hint: use (12), Sec. 7.3.
16. Let A be a given square matrix, and let X be the matrix whose column vectors x_1, \dots, x_n are eigenvectors of A corresponding to the eigenvalues $\lambda_1, \dots, \lambda_n$ of A . Show that

$$AX = (\lambda_1 x_1 \cdots \lambda_n x_n) = XD,$$

where D is the diagonal matrix with elements $\lambda_1, \dots, \lambda_n$ in the principal diagonal. (X is called the **modal matrix** and D the **spectral matrix** of A .)

17. If the x_i in Prob. 16 are linearly independent, show that $X^{-1}AX = D$.
18. The sum of the elements in the principal diagonal of a square matrix $C = (c_{ik})$ is called the **trace** of C . Thus

$$\text{trace } C = c_{11} + c_{22} + \cdots + c_{nn}.$$

Let $A = (a_{ik})$ and $B = (b_{ik})$ be two n -rowed square matrices. Show that

$$\text{trace } AB = \sum_{i=1}^n \sum_{l=1}^n a_{il} b_{li}.$$

19. If $\bar{A} = T^{-1}AT$, show that $\text{trace } \bar{A} = \text{trace } A$.
20. Show that the matrices A and $\bar{A} = T^{-1}AT$ have the same eigenvalues.
21. If x is an eigenvector of A corresponding to an eigenvalue λ , show that $y = T^{-1}x$ is an eigenvector of $\bar{A} = T^{-1}AT$ corresponding to the same eigenvalue λ .
22. Verify the statement of Prob. 20 for the matrices

$$A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad \bar{A} = T^{-1}AT \quad \text{where} \quad T = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}.$$

23. Find a real matrix which has real eigenvalues but is not symmetric. Does this contradict Theorem 3?

24. Find a real matrix which has purely imaginary eigenvalues but is not skew-symmetric. Does this contradict Theorem 3?
25. Show that the vertical vibrations of the mechanical system in Fig. 212 (masses of springs and damping neglected) are governed by the following system of differential equations:

$$\ddot{y}_1 = -ky_1 + k(y_2 - y_1)$$

$$\ddot{y}_2 = -k(y_2 - y_1) - ky_2$$

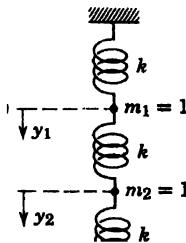


Fig. 212. Problem 25 (System in static equilibrium).

26. Show that the system in Prob. 25 may be written

$$\ddot{\mathbf{y}} = \mathbf{A}\mathbf{y}, \quad \text{where} \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \quad \text{and} \quad \mathbf{A} = \begin{pmatrix} -2k & k \\ k & -2k \end{pmatrix}.$$

27. To solve the vector equation in Prob. 26, substitute $\mathbf{y} = \mathbf{x}e^{\omega t}$. Show that this leads to the eigenvalue problem $\mathbf{Ax} = \lambda\mathbf{x}$ where $\lambda = \omega^2$. Find the eigenvalues and eigenvectors. Find the solution of the vector equation that satisfies the initial conditions $y_1(0) = 1$, $y_2(0) = 1$, $\dot{y}_1(0) = \sqrt{3k}$, $\dot{y}_2(0) = -\sqrt{3k}$.
28. Show that the vertical vibrations of the mechanical system in Fig. 213 are governed by the equation

$$\ddot{\mathbf{y}} = \mathbf{Ay}$$

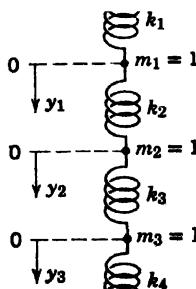


Fig. 213. Problem 28 (System in static equilibrium).

where

$$\mathbf{A} = \begin{pmatrix} -(k_1 + k_2) & k_2 & 0 \\ k_2 & -(k_2 + k_3) & k_3 \\ 0 & k_3 & -(k_3 + k_4) \end{pmatrix}$$

and substitution of $\mathbf{y} = \mathbf{x}e^{\omega t}$ yields $\mathbf{Ax} = \lambda\mathbf{x}$ where $\lambda = \omega^2$.

29. Let \mathbf{A} be the matrix in Prob. 28, with $k_1 = k_2 = k_3 = k_4 = k$. Show that the corresponding matrix $\mathbf{B} = \mathbf{A}/k + 2\mathbf{I}$ has the eigenvalues 0 , $+\sqrt{2}$, and $-\sqrt{2}$. Conclude from this that \mathbf{A} has the eigenvalues $-2k$, $(-2 + \sqrt{2})k$, and $(-2 - \sqrt{2})k$.
30. Verify that the coefficients of the characteristic equation

$$\lambda^3 + a\lambda^2 + b\lambda + c = 0,$$

corresponding to the matrix \mathbf{A} in Prob. 28, are positive, and conclude that the eigenvalues of \mathbf{A} must be negative real. What does this result mean with respect to the motion of the mechanical system in Fig. 213?

7.14 Bounds for Eigenvalues

To compute the eigenvalues of a given n -rowed square matrix we may proceed as follows. We first determine the characteristic polynomial, that is, we develop the characteristic determinant and then collect terms involving the same power of λ . In the second step we determine the roots of the characteristic polynomial; if n is large, this may be done by using an approximation method such as Newton's method. Obviously, for large n both steps require time-consuming computational operations. Therefore, it is desirable to develop more effective numerical methods which yield information about the eigenvalues of a given matrix. There are essentially two types of methods:

1. *Methods for obtaining bounds for eigenvalues,*
2. *Methods for computing approximate values for eigenvalues.*

We shall illustrate both types by some standard examples. For further methods see Ref. [D1] in Appendix 1.

The following interesting theorem by Gershgorin⁵ yields a region consisting of closed circular disks and containing all the eigenvalues of a given matrix.

Theorem 1. *Let λ be an eigenvalue of an arbitrary n -rowed square matrix $\mathbf{A} = (a_{jk})$. Then for some k ($1 \leq k \leq n$),*

$$(1) \quad |a_{kk} - \lambda| \leq |a_{k1}| + |a_{k2}| + \cdots + |a_{k,k-1}| + |a_{k,k+1}| + \cdots + |a_{kn}|.$$

Proof. Let \mathbf{x} be an eigenvector corresponding to that eigenvalue λ of \mathbf{A} . Then

$$(2) \quad \mathbf{Ax} = \lambda\mathbf{x} \quad \text{or} \quad (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}.$$

⁵ Bull. Acad. Sciences de l'URSS, Classe mathém., 7-e série, Leningrad, 1931, p. 749.

Let x_k be a component of \mathbf{x} which is the greatest in absolute value. Then $|x_m/x_k| \leq 1$ ($m = 1, \dots, n$). The vector equation (2) is equivalent to a system of n equations for the n components of the vectors on both sides, and the k th of these n equations is

$$a_{k1}x_1 + \cdots + a_{k,k-1}x_{k-1} + (a_{kk} - \lambda)x_k + a_{k,k+1}x_{k+1} + \cdots + a_{kn}x_n = 0.$$

From this we have

$$a_{kk} - \lambda = -a_{k1}\frac{x_1}{x_k} - \cdots - a_{k,k-1}\frac{x_{k-1}}{x_k} - a_{k,k+1}\frac{x_{k+1}}{x_k} - \cdots - a_{kn}\frac{x_n}{x_k}.$$

By taking absolute values on both sides of this equation, applying the triangle inequality

$$|a + b| \leq |a| + |b| \quad (a, b \text{ any complex numbers})$$

and observing that

$$\left| \frac{x_1}{x_k} \right| \leq 1, \quad \dots, \quad \left| \frac{x_n}{x_k} \right| \leq 1$$

we obtain (1), and the theorem is proved.

For each $k = 1, \dots, n$ the inequality (1) determines a closed circular disk in the complex λ -plane whose center is at a_{kk} and whose radius is given by the expression on the right-hand side of (1). Theorem 1 states that each of the eigenvalues of \mathbf{A} lies in one of these n disks.

Example 1. From Theorem 1 it follows that the eigenvalues of the matrix

$$\mathbf{A} = \begin{pmatrix} 26 & -2 & 2 \\ 2 & 21 & 4 \\ 4 & 2 & 28 \end{pmatrix}$$

lie in the three disks (Fig. 214)

D_1 : center at 26, radius $|-2| + 2 = 4$,

D_2 : center at 21, radius $2 + 4 = 6$,

and

D_3 : center at 28, radius $4 + 2 = 6$.

(The eigenvalues of \mathbf{A} are 30, 25, and 20.)

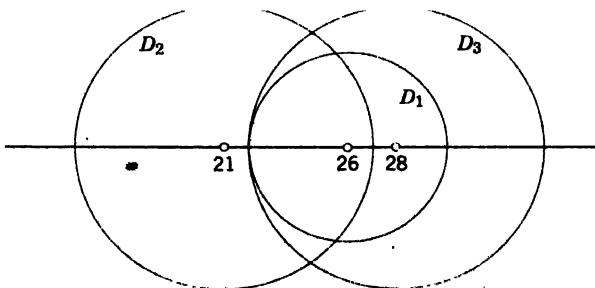


Fig. 214. Example 1.

Bounds for the absolute value of the eigenvalues result from the following theorem by Schur⁶ which we shall state without proof.

Theorem 2. Let $\mathbf{A} = (a_{jk})$ be an n -rowed square matrix, and let $\lambda_1, \dots, \lambda_n$ be its eigenvalues. Then

$$(3) \quad \sum_{i=1}^n |\lambda_i|^2 \leq \sum_{j=1}^n \sum_{k=1}^n |a_{jk}|^2 \quad (\text{Schur's inequality}).$$

In (3) the equality sign holds if, and only if, \mathbf{A} is such that

$$(4) \quad \bar{\mathbf{A}}^T \mathbf{A} = \mathbf{A} \bar{\mathbf{A}}^T.$$

Matrices which satisfy (4) are called **normal matrices**. It is not difficult to see that Hermitian, skew-Hermitian, and unitary matrices are normal, and so are real symmetric, skew-symmetric, and orthogonal matrices.

Let λ_m be any eigenvalue of the matrix \mathbf{A} in Theorem 2. Then $|\lambda_m|^2$ is less than or equal to the sum on the left side of (3), and by taking square roots we obtain from (3)

$$(5) \quad |\lambda_m| \leq \sqrt{\sum_{j=1}^n \sum_{k=1}^n |a_{jk}|^2}.$$

Example 2. For the matrix \mathbf{A} in Ex. 1 we obtain from (5)

$$|\lambda| \leq \sqrt{1949} < 44.2.$$

(The eigenvalues of \mathbf{A} are 30, 25, and 20; thus $30^2 + 25^2 + 20^2 = 1925 < 1949$; in fact, \mathbf{A} is not normal.)

The preceding theorems are valid for any real or complex square matrix. There are other theorems which hold for special classes of matrices only. The following theorem by Frobenius,⁷ which we state without proof, is of this type.

Theorem 3. Let \mathbf{A} be a real square matrix whose elements are all positive. Then \mathbf{A} has at least one real positive eigenvalue λ , and the corresponding eigenvector can be chosen real and such that all its components are positive.

From this theorem we may derive the following useful result by Collatz.⁸

Theorem 4. Let $\mathbf{A} = (a_{jk})$ be a real n -rowed square matrix whose elements are all positive. Let \mathbf{x} be any real vector whose components x_1, \dots, x_n are all positive, and let y_1, \dots, y_n be the components of the vector $\mathbf{y} = \mathbf{Ax}$. Then the closed interval on the real axis bounded by the smallest and the largest of the n quotients $q_j = y_j/x_j$ contains at least one eigenvalue of \mathbf{A} .

Proof. We have $\mathbf{Ax} = \mathbf{y}$ or

$$(6) \quad \mathbf{y} - \mathbf{Ax} = \mathbf{0}.$$

⁶ *Mathematische Annalen*, Vol. 66, 1909, p. 488.

⁷ *Sitzungsberichte Preuss. Akad. Wiss. Math-phys. Klasse*, Berlin, 1908, p. 471.

⁸ L. Collatz, *Eigenwertaufgaben mit technischen Anwendungen*. Leipzig: Akademische Verlagsgesellschaft, 1949, p. 291.

The transpose \mathbf{A}^T satisfies the conditions of Theorem 3. Hence \mathbf{A}^T has a positive eigenvalue λ and, corresponding to this eigenvalue, an eigenvector \mathbf{u} whose components u_j are all positive. Thus $\mathbf{A}^T\mathbf{u} = \lambda\mathbf{u}$, and by taking the transpose, $\mathbf{u}^T\mathbf{A} = \lambda\mathbf{u}^T$. From this and (6),

$$\mathbf{u}^T(\mathbf{y} - \mathbf{Ax}) = \mathbf{u}^T\mathbf{y} - \mathbf{u}^T\mathbf{Ax} = \mathbf{u}^T(\mathbf{y} - \lambda\mathbf{x}) = 0$$

or at length

$$\sum_{j=1}^n u_j(y_j - \lambda x_j) = 0.$$

Since all the components u_j are positive, it follows that

$$(7) \quad \begin{aligned} y_j - \lambda x_j &\geq 0, & \text{that is, } q_j &\geq \lambda & \text{for at least one } j, \\ y_j - \lambda x_j &\leq 0, & \text{that is, } q_j &\leq \lambda & \text{for at least one } j. \end{aligned} \quad \text{and}$$

Since \mathbf{A} and \mathbf{A}^T have the same eigenvalues, λ is an eigenvalue of \mathbf{A} , and from (7) the statement of the theorem follows.

Example 3. Let

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 1 \end{pmatrix}.$$

Choosing

$$\mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \text{we obtain} \quad \mathbf{y} = \mathbf{Ax} = \begin{pmatrix} 6 \\ 12 \\ 10 \end{pmatrix}.$$

Hence $q_1 = 6$, $q_2 = 12$, $q_3 = 10$. From Theorem 4 it follows that one of the eigenvalues of \mathbf{A} must lie in the interval $6 \leq \lambda \leq 12$.

Of course the length of such an interval depends on the choice of \mathbf{x} . Choosing

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad \text{we obtain} \quad \mathbf{y} = \mathbf{Ax} = \begin{pmatrix} 11 \\ 22 \\ 17 \end{pmatrix}.$$

$q_1 = 11$, $q_2 = 11$, $q_3 = 8.5$, and therefore the shorter interval $8.5 \leq \lambda \leq 11$ which must contain an eigenvalue of \mathbf{A} . The student may show that $\lambda = 10$ is an eigenvalue of \mathbf{A} .

Problems

Using Theorem 1, determine and graph disks which contain the eigenvalues of the following matrices.

1. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

2. $\begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}$

3. $\begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix}$

4. $\begin{pmatrix} 6 & 0 & -3 \\ 0 & 6 & 3 \\ -3 & 3 & 2 \end{pmatrix}$

5. $\begin{pmatrix} -9 & 1 & 0 \\ 1 & -9 & 1 \\ 0 & 1 & -9 \end{pmatrix}$

6. $\begin{pmatrix} 0 & 0 & 3i \\ 0 & 2-i & 1+i \\ 1+2i & 0 & 0 \end{pmatrix}$

7. Show that the matrix in Ex. 1 is not normal.
 8. Show that the matrix in Prob. 1 is not normal and find its eigenvalues.
 9. Verify that the matrix in Prob. 2 is normal, and for this matrix the equality sign holds in (3).
 10. Show that Hermitian, skew-Hermitian, and unitary matrices are normal.
 11. Show that if A is normal and U is unitary and both matrices are n -rowed, then $B = U^{-1}AU$ is normal.
- Using (5), obtain an upper bound for the absolute value of the eigenvalues of the matrix:
12. In Prob. 2
 13. In Prob. 3
 14. In Prob. 4
 15. In Ex. 3
 16. In Prob. 5
 17. In Prob. 6
 18. Verify Theorem 3 for the matrix in Prob. 3.

Apply Theorem 4 to the following matrices, choosing the given vectors as vectors \mathbf{x} .

$$19. \begin{pmatrix} 17 & 8 & 1 \\ 8 & 18 & 8 \\ 1 & 8 & 17 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix} \quad 20. \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

7.15 Determination of Eigenvalues by Iteration

A standard method for computing approximate values of the eigenvalues of matrices is the so-called *iteration method*. We shall first explain the practical procedure, then illustrate it by an example, and finally consider the theoretical background.

Let $A = (a_{jk})$ be a given n -rowed real symmetric matrix. Take an arbitrary real vector \mathbf{x}_0 ($\neq 0$) with n components and compute the vectors

$$(1) \quad \mathbf{x}_1 = A\mathbf{x}_0, \quad \mathbf{x}_2 = A\mathbf{x}_1, \quad \dots, \quad \mathbf{x}_s = A\mathbf{x}_{s-1}.$$

For simplifying notation, denote \mathbf{x}_{s-1} by \mathbf{x} and \mathbf{x}_s by \mathbf{y} so that

$$(2) \quad \mathbf{y} = A\mathbf{x}.$$

Compute the scalar products

$$m_0 = \mathbf{x}^T \mathbf{x} \quad \text{and} \quad m_1 = \mathbf{x}^T \mathbf{y}.$$

Then the quotient

$$(3) \quad q = \frac{m_1}{m_0}$$

is an approximation for one of the eigenvalues of A (in general, for that eigenvalue which has the greatest absolute value among all eigenvalues of A).

Example 1. Let

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 1 \end{pmatrix} \quad \text{and choose} \quad \mathbf{x}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Then

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 14 \\ 28 \\ 18 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} 124 \\ 248 \\ 228 \end{pmatrix}, \quad \mathbf{x}_4 = \begin{pmatrix} 1304 \\ 2608 \\ 2088 \end{pmatrix}$$

Taking $\mathbf{x} = \mathbf{x}_3$ and $\mathbf{y} = \mathbf{x}_4$, we have

$$m_0 = \mathbf{x}^T \mathbf{x} = 128\ 864, \quad m_1 = \mathbf{x}^T \mathbf{y} = 1\ 284\ 544, \quad q \approx 9.968.$$

The reader may show that $\lambda = 10$ is an eigenvalue of \mathbf{A} .

Taking $\mathbf{x} = \mathbf{x}_2$ and $\mathbf{y} = \mathbf{x}_3$, we obtain

$$m_0 = 1304, \quad m_1 = 12\ 784, \quad q \approx 9.80,$$

a less accurate approximation for $\lambda = 10$. In general, the accuracy of q will increase with increasing s ; that is, the more vectors (1) we compute, the more accurate q will be. However, by looking at the theoretical basis of the method, we shall now see that this statement is not always true.

The eigenvectors $\mathbf{z}_1, \dots, \mathbf{z}_n$ of a real symmetric matrix $\mathbf{A} = (a_{jk})$ can be chosen as real unit vectors and such that they are mutually perpendicular:

$$(4) \quad \mathbf{z}_j^T \mathbf{z}_j = 1, \quad \mathbf{z}_j^T \mathbf{z}_k = 0 \quad (j \neq k)$$

(proof in Ref. [D7]). Hence our arbitrary vector \mathbf{x}_0 can be expressed as a linear combination of the eigenvectors, say,

$$(5) \quad \mathbf{x}_0 = c_1 \mathbf{z}_1 + \dots + c_n \mathbf{z}_n.$$

It follows that \mathbf{x}_1 in (1) then has the representation

$$\mathbf{x}_1 = \mathbf{A} \mathbf{x}_0 = c_1 \mathbf{A} \mathbf{z}_1 + \dots + c_n \mathbf{A} \mathbf{z}_n.$$

Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues corresponding to $\mathbf{z}_1, \dots, \mathbf{z}_n$, respectively. Then, by definition

$$\mathbf{A} \mathbf{z}_1 = \lambda_1 \mathbf{z}_1, \dots, \mathbf{A} \mathbf{z}_n = \lambda_n \mathbf{z}_n,$$

and the representation for \mathbf{x}_1 becomes

$$\mathbf{x}_1 = c_1 \lambda_1 \mathbf{z}_1 + \dots + c_n \lambda_n \mathbf{z}_n.$$

In the same way it follows from this that

$$\mathbf{x}_2 = \mathbf{A} \mathbf{x}_1 = c_1 \lambda_1^2 \mathbf{z}_1 + \dots + c_n \lambda_n^2 \mathbf{z}_n,$$

and for the last two vectors in (1) we similarly obtain

$$(6) \quad \begin{aligned} \mathbf{x} &\equiv \mathbf{x}_{s-1} = c_1 \lambda_1^{s-1} \mathbf{z}_1 + \dots + c_n \lambda_n^{s-1} \mathbf{z}_n \\ \mathbf{y} &\equiv \mathbf{x}_s = c_1 \lambda_1^s \mathbf{z}_1 + \dots + c_n \lambda_n^s \mathbf{z}_n. \end{aligned}$$

We now form the scalar products m_0 and m_1 ; because of the orthogonality (4), we obtain from (6)

$$\begin{aligned} m_0 &= \mathbf{x}^T \mathbf{x} = c_1^2 \lambda_1^{2s-2} + \dots + c_n^2 \lambda_n^{2s-2}, \\ m_1 &= \mathbf{x}^T \mathbf{y} = c_1^2 \lambda_1^{2s-1} + \dots + c_n^2 \lambda_n^{2s-1}. \end{aligned}$$

From these expressions it follows that

$$q = \frac{m_1}{m_0} = \frac{c_1^2 \lambda_1^{2s-1} + \dots}{c_1^2 \lambda_1^{2s-2} + \dots} \approx \lambda_1$$

provided the terms indicated by the dots are small compared with the two explicitly written terms. If $c_1 \neq 0$, λ_1 is the greatest eigenvalue in absolute value, and the other eigenvalues are small in absolute value, then q will be a good approximation of λ_1 and will become even more accurate for larger s . Now from (5) and (4),

$$\mathbf{z}_1^T \mathbf{x}_0 = c_1 \mathbf{z}_1^T \mathbf{z}_1 = c_1$$

and we see that $c_1 \neq 0$, if \mathbf{x}_0 and \mathbf{z}_1 are not orthogonal. We do not know \mathbf{z}_1 , but the chance of choosing for \mathbf{x}_0 just a vector which is perpendicular to \mathbf{z}_1 will be small. A more serious handicap of the method is that another eigenvalue of \mathbf{A} may differ very little from λ_1 in absolute value, and in this case q will be an inaccurate approximation even for large s . The reader may consider the matrix

$$(7) \quad \mathbf{A} = \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}, \quad \text{choose} \quad \mathbf{x}_0 = \begin{pmatrix} 3 \\ -1 \end{pmatrix},$$

and find $q = 0$ for all s , the eigenvalues being 5 and -5 . Therefore, to use the iteration method effectively we need bounds for the error of the approximate value q . Such a bound is given by the following theorem.

Theorem 1. *Let \mathbf{A} be an n -rowed real symmetric matrix. Let \mathbf{x} ($\neq 0$) be any real vector with n components. Furthermore, let*

$$\mathbf{y} = \mathbf{Ax}, \quad m_0 = \mathbf{x}^T \mathbf{x}, \quad m_1 = \mathbf{x}^T \mathbf{y}, \quad m_2 = \mathbf{y}^T \mathbf{y}.$$

Then the quotient

$$q = \frac{m_1}{m_0}$$

is an approximation for an eigenvalue λ of \mathbf{A} , and if we set $\epsilon = \lambda + \epsilon$, so that ϵ is the error of q , then

$$(8) \quad |\epsilon| \leq \sqrt{\frac{m_2}{m_0} - q^2}.$$

Proof. Let δ^2 denote the radicand in (8). Then, since $m_1 = qm_0$,

$$(9) \quad (\mathbf{y} - q\mathbf{x})^T(\mathbf{y} - q\mathbf{x}) = m_2 - 2qm_1 + q^2m_0 = m_2 - q^2m_0 = \delta^2m_0.$$

Let $\mathbf{z}_1, \dots, \mathbf{z}_n$ be real orthogonal unit eigenvectors of \mathbf{A} and let

$$\mathbf{x} = a_1 \mathbf{z}_1 + \dots + a_n \mathbf{z}_n.$$

Then, as before,

$$\mathbf{y} = a_1 \lambda_1 \mathbf{z}_1 + \dots + a_n \lambda_n \mathbf{z}_n$$

and

$$(10) \quad m_0 = \mathbf{x}^T \mathbf{x} = a_1^2 + \dots + a_n^2$$

Hence in (9),

$$\mathbf{y} - q\mathbf{x} = a_1(\lambda_1 - q)\mathbf{z}_1 + \cdots + a_n(\lambda_n - q)\mathbf{z}_n.$$

By using the orthogonality of the \mathbf{z} , we thus obtain from (9)

$$\delta^2 m_0 = a_1^2(\lambda_1 - q)^2 + \cdots + a_n^2(\lambda_n - q)^2.$$

Replacing each $(\lambda_i - q)^2$ by the smallest of these terms and using (10), we have

$$\delta^2 m_0 \geq (\lambda_c - q)^2(a_1^2 + \cdots + a_n^2) = (\lambda_c - q)^2 m_0$$

where λ_c is the eigenvalue to which q is closest. From this, (8) follows, and the theorem is proved.

Example 2. From (8) we obtain the following bound for the error ϵ of the value $q \approx 9.968$ in Ex. 1:

$$|\epsilon| \leq \sqrt{\frac{12861.824}{128.64} - 9.968^2} \approx \sqrt{0.45} \approx 0.67.$$

This shows that the exact value of the eigenvalue must lie between 9.29 and 10.64.

Problems

1. Let

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 3 & 2 \\ 1 & 2 & 3 \end{pmatrix} \quad \text{and choose} \quad \mathbf{x}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Compute \mathbf{x}_1 , \mathbf{x}_2 , \mathbf{x}_3 . Take \mathbf{x}_2 for \mathbf{x} and \mathbf{x}_3 for \mathbf{y} . Show that q deviates from 5 (the exact value of the largest eigenvalue of \mathbf{A}) by 1.7 per cent, approximately. Find a bound for the error of q from (8).

Let \mathbf{A} be as in Prob. 1. Compute \mathbf{x}_1 , the corresponding q , and a bound for the error, choosing

$$2. \quad \mathbf{x}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$3. \quad \mathbf{x}_0 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$4. \quad \mathbf{x}_0 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

5. Show that the eigenvalues of the matrix \mathbf{A} in Prob. 1 are 0, 3, 5, and indicate which eigenvalue is approximated by the results of Probs. 2-4.

Choosing for \mathbf{x}_0 the column vector with components 1, 1, 1, 1, compute \mathbf{x}_1 , \mathbf{x}_2 and approximations $q = \mathbf{x}_1^T \mathbf{x}_0 / \mathbf{x}_0^T \mathbf{x}_0$, $q = \mathbf{x}_2^T \mathbf{x}_1 / \mathbf{x}_1^T \mathbf{x}_1$ and corresponding error bounds for an eigenvalue of each of the following symmetric matrices.

$$6. \quad \begin{vmatrix} 3 & 2 & 0 & 1 \\ 2 & 0 & 5 & -1 \\ 0 & 5 & 2 & 1 \\ 1 & -1 & 1 & 4 \end{vmatrix} \quad 7. \quad \begin{vmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & -1 & 0 \\ 0 & -1 & 3 & 0 \\ 0 & -1 & 0 & 1 \end{vmatrix} \quad 8. \quad \begin{vmatrix} 2 & 0 & 1 & 0 \\ 0 & 0 & 3 & 1 \\ 1 & 3 & 4 & -2 \\ 0 & 1 & -2 & 0 \end{vmatrix}$$

9. Show that if x and y are eigenvectors, then $\epsilon = 0$ in (8).
10. Determine the eigenvalues and eigenvectors of the matrix (7) and explain the reason why $q = 0$ for all s . Choose another vector x_0 and repeat the iteration. Note that in (7), $x_2 = 25x_0$ and conclude that 25 must be an eigenvalue of A^2 . What follows for A from this result?

Chapter 8

Fourier Series and Integrals

Periodic functions occur frequently in engineering problems. Their representation in terms of simple periodic functions, such as sine and cosine, is a matter of great practical importance, which leads to Fourier series. These series, named after the French physicist JOSEPH FOURIER (1768–1830), represent a very powerful tool in connection with various problems involving ordinary and partial differential equations.

In the present chapter the basic concepts, facts, and techniques in connection with Fourier series will be considered and illustrated by various examples. Some important engineering applications will be included; further applications will occur in the following chapter on partial differential equations and boundary value problems.

The *theory* of Fourier series is rather complicated, but the *application* of these series is simple. Fourier series are, in a certain sense, more universal than Taylor series, because many *discontinuous* periodic functions of practical interest can be developed in Fourier series, but, of course, do not have Taylor series representations.

Prerequisite for this chapter: elementary integral calculus.

Sections which may be omitted in a shorter course: 8.8, 8.9, 8.11–8.13.

References: Appendix 1, Part E.

Answers to problems: Appendix 2.

8.1 Periodic Functions. Trigonometric Series

A function $f(x)$ is said to be **periodic** if it is defined for all real x and if there is some positive number T such that

$$(1) \quad f(x + T) = f(x) \quad \text{for all } x.$$

The number T is then called a **period¹** of $f(x)$. The graph of such a function

¹ The smallest positive period T of a function $f(x)$, not a constant, is often called the *primitive period* of $f(x)$. For example, the primitive periods of $\sin x$ and $\sin 2x$ are 2π and π , respectively.

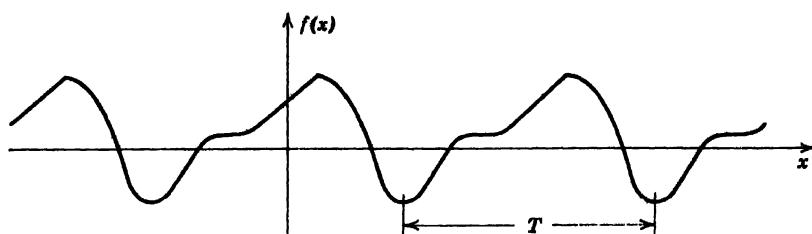


Fig. 215. Periodic function.

is obtained by periodic repetition of its graph in any interval of length T (Fig. 215).

From (1) it follows that, if n is any integer,

$$f(x + nT) = f(x) \quad \text{for all } x,$$

so that any integral multiple nT ($n \neq 0$) of T is also a period. Furthermore, if $f(x)$ and $g(x)$ have the period T , then the function

$$h(x) = af(x) + bg(x) \quad (a, b \text{ constant})$$

has the period T .

Familiar examples of periodic functions are the sine and cosine functions, and we note that the function $f = c = \text{const}$ is also a periodic function in the sense of the definition, because it satisfies (1) for any positive T .

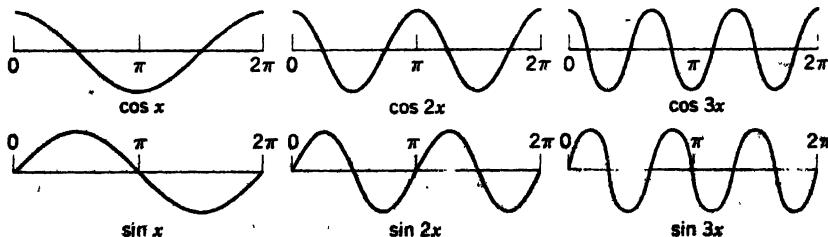
Our problem in the first few sections of this chapter will be the representation of various functions of period 2π in terms of the simple functions

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \dots, \cos nx, \sin nx, \dots$$

which have the period 2π (Fig. 216). The series which will arise in this connection will be of the form

$$(2) \quad a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x +$$

where $a_0, a_1, a_2, \dots, b_1, b_2, \dots$ are real constants. Such a series is called a **trigonometric series**, and the a_n and b_n are called the **coefficients** of the series. We see that each term of the series has the period 2π . Hence if the series converges, its sum will be a function of period 2π .

Fig. 216. Cosine and sine functions having the period 2π .

Periodic functions that occur in engineering problems are often rather complicated, and it is therefore desirable to represent these functions in terms of simple periodic functions. We shall see that almost any periodic function $f(x)$ of period 2π that appears in applications, for example, in connection with vibrations, can be represented by a trigonometric series, and we shall derive formulas for the coefficients in (2) in terms of $f(x)$ such that (2) converges and has the sum $f(x)$. Later we shall extend our results to functions of arbitrary period; this extension will turn out to be quite simple.

Problems

Find the smallest positive period T of the following functions.

1. $\cos x, \sin x, \cos 2x, \sin 2x, \cos nx, \sin nx$
2. $\cos 2\pi x, \sin 2\pi x, \cos(2\pi x/k), \sin(2\pi x/k), \cos(2\pi nx/k), \sin(2\pi nx/k)$

Plot accurate graphs of the following functions.

$$3. \sin x, \quad \sin x + \frac{1}{3} \sin 3x, \quad \sin x + \frac{1}{3} \sin 3x + \frac{1}{5} \sin 5x,$$

$$f(x) = \begin{cases} -\pi/4 & \text{when } -\pi < x < 0 \\ \pi/4 & \text{when } 0 < x < \pi \end{cases} \quad \text{and } f(x + 2\pi) = f(x)$$

$$4. \sin 2\pi x, \quad \sin 2\pi x + \frac{1}{3} \sin 6\pi x, \quad \sin 2\pi x + \frac{1}{3} \sin 6\pi x + \frac{1}{5} \sin 10\pi x$$

$$5. \frac{\pi}{4} + \cos x, \quad \frac{\pi}{4} + \cos x - \frac{1}{3} \cos 3x, \quad \frac{\pi}{4} + \cos x - \frac{1}{3} \cos 3x + \frac{1}{5} \cos 5x$$

$$6. f(x) = x \text{ when } -\pi < x < \pi \text{ and } f(x + 2\pi) = f(x)$$

$$7. f(x) = x^2 \text{ when } -\pi < x < \pi \text{ and } f(x + 2\pi) = f(x)$$

$$8. f(x) = x^3 \text{ when } -\pi < x < \pi \text{ and } f(x + 2\pi) = f(x)$$

$$9. f(x) = e^x \text{ when } -\pi < x < \pi \text{ and } f(x + 2\pi) = f(x)$$

$$10. f(x) = \begin{cases} 0 & \text{when } -\pi < x < 0 \\ x & \text{when } 0 < x < \pi \end{cases} \quad \text{and } f(x + 2\pi) = f(x)$$

$$11. f(x) = \begin{cases} -x^2 & \text{when } -\pi < x < 0 \\ x^2 & \text{when } 0 < x < \pi \end{cases} \quad \text{and } f(x + 2\pi) = f(x)$$

$$12. f(x) = \begin{cases} 1 & \text{when } -\pi < x < 0 \\ \cos^2 2x & \text{when } 0 < x < \pi \end{cases} \quad \text{and } f(x + 2\pi) = f(x)$$

13. If T is a period of $f(x)$, show that $nT, n = \pm 2, \pm 3, \dots$ is a period of that function.

14. If $f(x)$ and $g(x)$ have the period T , show that $h = af + bg$ (a, b constant) has the period T .

15. Show that the function $f(x) = \text{const}$ is a periodic function of period T for any value of T .

16. If $f(x)$ is a periodic function of x of period T , show that $f(ax)$, $a \neq 0$, is a periodic function of x of period T/a , and $f(x/b)$, $b \neq 0$, is a periodic function of x of period bT . Verify these results for $f(x) = \sin x$, $a = b = 2$.

Evaluate the following integrals where $n = 0, 1, 2, \dots$. (These are typical examples of integrals which will be needed in our further work.)

17. $\int_0^{\pi/2} \cos nx \, dx$

18. $\int_0^\pi \sin nx \, dx$

19. $\int_{-\pi/2}^{\pi/2} x \cos nx \, dx$

20. $\int_{-\pi/2}^{\pi/2} x \sin nx \, dx$

21. $\int_{-\pi}^\pi x \sin nx \, dx$

22. $\int_0^\pi x \sin nx \, dx$

23. $\int_0^\pi x^2 \cos nx \, dx$

24. $\int_0^\pi e^x \cos nx \, dx$

25. $\int_{-\pi}^0 e^x \sin nx \, dx$

8.2 Fourier Series. Euler's Formulas

Let us suppose that $f(x)$ is a periodic function with period 2π which can be represented by a trigonometric series

$$(1) \quad f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx).$$

Given such a function $f(x)$, we want to determine the coefficients a_n and b_n in the corresponding series (1).

We first determine a_0 . Integrating on both sides of (1) from $-\pi$ to π , we have

$$\int_{-\pi}^\pi f(x) \, dx = \int_{-\pi}^\pi \left[a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \right] \, dx.$$

If term-by-term integration of the series is allowed,² then we obtain

$$\int_{-\pi}^\pi f(x) \, dx = a_0 \int_{-\pi}^\pi \, dx + \sum_{n=1}^{\infty} \left(a_n \int_{-\pi}^\pi \cos nx \, dx + b_n \int_{-\pi}^\pi \sin nx \, dx \right).$$

The first term on the right equals $2\pi a_0$, while all the other integrals are zero, as can be readily seen by performing the integrations. Hence our first result is

$$(2) \quad a_0 = \frac{1}{2\pi} \int_{-\pi}^\pi f(x) \, dx,$$

the area under the curve of $f(x)$ from $-\pi$ to π , divided by 2π .

We now determine a_1, a_2, \dots by a similar procedure. We multiply (1) by $\cos mx$, where m is any fixed positive integer, and then integrate from $-\pi$ to π , finding

$$(3) \quad \int_{-\pi}^\pi f(x) \cos mx \, dx = \int_{-\pi}^\pi \left[a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \right] \cos mx \, dx.$$

² This is justified, for instance, in the case of uniform convergence (cf. Theorem 3 in Sec. 14.4).

By term-by-term integration the right-hand side becomes

$$a_0 \int_{-\pi}^{\pi} \cos mx dx + \sum_{n=1}^{\infty} \left[a_n \int_{-\pi}^{\pi} \cos nx \cos mx dx + b_n \int_{-\pi}^{\pi} \sin nx \cos mx dx \right].$$

The first integral is zero, and so is the last integral, because its integrand is an odd function (Sec. 0.1). By applying (11), Sec. 0.1, to the second integral we obtain

$$\int_{-\pi}^{\pi} \cos nx \cos mx dx = \frac{1}{2} \int_{-\pi}^{\pi} \cos(n+m)x dx + \frac{1}{2} \int_{-\pi}^{\pi} \cos(n-m)x dx.$$

In this formula the first integral on the right is 0 for all m and n under consideration; the last integral is zero when $n \neq m$ and is π when $n = m$. Hence the right-hand side in (3) equals $a_m \pi$, and our second result is

$$(4) \quad a_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos mx dx, \quad m = 1, 2, \dots$$

We finally determine b_1, b_2, \dots in (1). If we multiply (1) by $\sin mx$, where m is any fixed positive integer, and then integrate from $-\pi$ to π , we have

$$(5) \quad \int_{-\pi}^{\pi} f(x) \sin mx dx = \int_{-\pi}^{\pi} \left[a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \right] \sin mx dx.$$

Integrating term-by-term, we see that the right-hand side becomes

$$a_0 \int_{-\pi}^{\pi} \sin mx dx + \sum_{n=1}^{\infty} \left[a_n \int_{-\pi}^{\pi} \cos nx \sin mx dx + b_n \int_{-\pi}^{\pi} \sin nx \sin mx dx \right].$$

The first integral is zero. The next integral is of the type considered before, and we know that it is zero for all $n = 1, 2, \dots$. The last integral can be transformed by means of (11), Sec. 0.1; we find

$$\int_{-\pi}^{\pi} \sin nx \sin mx dx = \frac{1}{2} \int_{-\pi}^{\pi} \cos(n-m)x dx - \frac{1}{2} \int_{-\pi}^{\pi} \cos(n+m)x dx.$$

The last term is zero. The first term on the right is zero when $n \neq m$ and is π when $n = m$. Since in (5) this term is multiplied by b_m , the right-hand side in (5) is equal to $b_m \pi$, and our last result is

$$b_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin mx dx, \quad m = 1, 2,$$

Writing n in place of m in this formula and in (4), we altogether have the so-called Euler formulas:

$$(6) \quad \begin{aligned} (a) \quad a_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx \\ (b) \quad a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx \\ (c) \quad b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx \end{aligned} \quad n = 1, 2, \dots$$

A periodic function $f(x)$ with period 2π being given, we may compute the a_n and b_n by (6) and form the trigonometric series

$$(7) \quad a_0 + a_1 \cos x + b_1 \sin x + \cdots + a_n \cos nx + b_n \sin nx + \cdots.$$

This series is then called the **Fourier series** corresponding to $f(x)$, and its coefficients obtained from (6) are called the **Fourier coefficients** of $f(x)$.

Note that because of the periodicity of the integrands the interval of integration in (6) may be replaced by any other interval of length 2π , for instance, by the interval $0 \leq x \leq 2\pi$.

From the definition of a definite integral it follows that, if $f(x)$ is continuous or merely piecewise continuous (continuous except for finitely many finite jumps in the interval of integration), the integrals in (6) exist and we may compute the Fourier coefficients of $f(x)$ by (6). The remaining question, whether the Fourier series thus obtained converges and has the sum $f(x)$, will be considered later in this section.

Let us illustrate the practical use of (6) by a simple example. Numerous other examples will occur in the following sections.

Example 1. Find the Fourier coefficients of the periodic function $f(x)$ in Fig. 217a. The analytic representation is

$$f(x) = \begin{cases} -k & \text{when } -\pi < x < 0 \\ k & \text{when } 0 < x < \pi \end{cases} \quad \text{and} \quad f(x + 2\pi) = f(x).$$

Functions of this type may occur as external forces acting on mechanical systems, electromotive forces in electric circuits, etc.

The area under the curve of $f(x)$ between $-\pi$ and π is zero and, therefore, by (6a), $a_0 = 0$. From (6b),

$$\begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx = \frac{1}{\pi} \left[\int_{-\pi}^0 (-k) \cos nx dx + \int_0^{\pi} k \cos nx dx \right] \\ &= \frac{1}{\pi} \left[-k \frac{\sin nx}{n} \Big|_{-\pi}^0 + k \frac{\sin nx}{n} \Big|_0^{\pi} \right] = 0 \end{aligned}$$

because $\sin nx = 0$ at $-\pi$, 0, and π for all $n = 1, 2, \dots$. From (6c),

$$\begin{aligned} b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx = \frac{1}{\pi} \left[\int_{-\pi}^0 (-k) \sin nx dx + \int_0^{\pi} k \sin nx dx \right] \\ &= \frac{1}{\pi} \left[k \frac{\cos nx}{n} \Big|_{-\pi}^0 - k \frac{\cos nx}{n} \Big|_0^{\pi} \right]. \end{aligned}$$

Since $\cos(-\alpha) = \cos \alpha$ and $\cos 0 = 1$, this yields

$$b_n = \frac{k}{n\pi} [\cos 0 - \cos(-n\pi) - \cos n\pi + \cos 0] = \frac{2k}{n\pi} (1 - \cos n\pi).$$

Now, $\cos \pi = -1$, $\cos 2\pi = 1$, $\cos 3\pi = -1$ etc., in general,

$$\cos n\pi = \begin{cases} -1 & \text{for odd } n, \\ 1 & \text{for even } n, \end{cases} \quad \text{and thus} \quad 1 - \cos n\pi = \begin{cases} 2 & \text{for odd } n, \\ 0 & \text{for even } n. \end{cases}$$

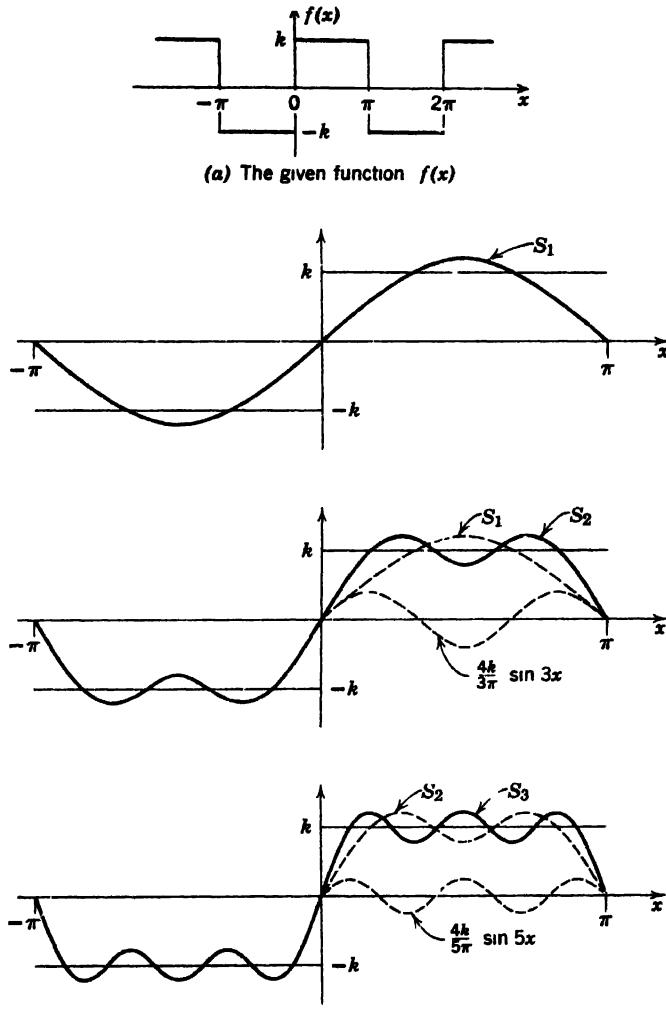


Fig. 217. Example 1.

Hence the Fourier coefficients b_n of our function are

$$b_1 = \frac{4k}{\pi}, \quad b_2 = 0, \quad b_3 = \frac{4k}{3\pi}, \quad b_4 = 0, \quad b_5 = \frac{4k}{5\pi}, \dots,$$

and since the a_n are zero, the corresponding Fourier series is

$$\frac{4k}{\pi} \left(\sin x + \frac{1}{3} \sin 3x + \frac{1}{5} \sin 5x + \dots \right).$$

The partial sums are

$$S_1 = \frac{4k}{\pi} \sin x, \quad S_2 = \frac{4k}{\pi} \left(\sin x + \frac{1}{3} \sin 3x \right), \quad \text{etc.}$$

and their graphs in Fig. 217 seem to indicate that the series is convergent and has the sum $f(x)$, the given function. We notice that at $x = 0$ and $x = \pi$, the points of discontinuity of $f(x)$, all partial sums have the value zero, the arithmetic mean of the values $-k$ and k of our function.

Furthermore, assuming that $f(x)$ is the sum of the series and setting $x = \pi/2$, we have

$$f\left(\frac{\pi}{2}\right) = k = \frac{4k}{\pi} \left(1 - \frac{1}{3} + \frac{1}{5} - + \cdots\right)$$

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + - \cdots = \frac{\pi}{4}.$$

This illustrates that the values of various series with constant terms can be obtained by evaluating Fourier series at specific points.

The class of functions which can be represented by Fourier series is surprisingly large and general. Corresponding sufficient conditions covering almost any conceivable engineering application are as follows.

Theorem 1. If a periodic function $f(x)$ with period 2π is piecewise continuous³ in the interval $-\pi \leq x \leq \pi$ and has a left- and right-hand derivative⁴ at each point of that interval, then the corresponding Fourier series (7) [with coefficients (6)] is convergent. Its sum is $f(x)$, except at a point x_0 at which $f(x)$ is discontinuous and the sum of the series is the average of the left- and right-hand limits⁴ of $f(x)$ at x_0 .

Remark. If the Fourier series corresponding to a function $f(x)$ converges with the sum $f(x)$ as characterized in Theorem 1, the series will be called the Fourier series of $f(x)$, we write

$$f(x) = a_0 + a_1 \cos x + b_1 \sin x + \cdots + a_n \cos nx + b_n \sin nx + \cdots,$$

and we say that $f(x)$ is represented by this Fourier series. Since insertion of parentheses in a convergent series yields a new convergent series having the

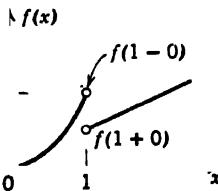


Fig. 218. Left- and right-hand limits

$$f(1-0) = 1,$$

$$f(1+0) = \frac{1}{2}$$

of the function

$$f(x) = \begin{cases} x^2 & \text{when } x < 1 \\ x/2 & \text{when } x > 1. \end{cases}$$

³ Definition in Sec. 4.1.

⁴ The left-hand limit of $f(x)$ at x_0 is defined as the limit of $f(x)$ as x approaches x_0 from the left and is frequently denoted by $f(x_0 - 0)$. Thus

$f(x_0 - 0) = \lim_{h \rightarrow 0} f(x_0 - h)$ as $h \rightarrow 0$ through positive values.

The right-hand limit is denoted by $f(x_0 + 0)$ and

$f(x_0 + 0) = \lim_{h \rightarrow 0} f(x_0 + h)$ as $h \rightarrow 0$ through positive values.

The left- and right-hand derivatives of $f(x)$ at x_0 are defined as the limits of

$\frac{f(x_0 - h) - f(x_0 - 0)}{-h}$ and $\frac{f(x_0 + h) - f(x_0 + 0)}{h}$,

respectively, as $h \rightarrow 0$ through positive values. Of course if $f(x)$ is continuous at x_0 the last term in both numerators is simply $f(x_0)$.

same sum as the original series (proof in Sec. 13.4), we may write more briefly

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx).$$

Proof of convergence in Theorem 1 for a continuous function $f(x)$ having continuous first and second derivatives. Integrating (6b) by parts, we obtain

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx = \frac{f(x) \sin nx}{n\pi} \Big|_{-\pi}^{\pi} - \frac{1}{n\pi} \int_{-\pi}^{\pi} f'(x) \sin nx dx.$$

The first term on the right is zero. A second integration by parts gives

$$a_n = \frac{f'(x) \cos nx}{n^2\pi} \Big|_{-\pi}^{\pi} - \frac{1}{n^2\pi} \int_{-\pi}^{\pi} f''(x) \cos nx dx.$$

The first term on the right is zero because of the periodicity and continuity of $f'(x)$. Since f'' is continuous in the interval of integration,

$$|f''(x)| < M$$

for an appropriate constant M . Furthermore, $|\cos nx| \leq 1$. It follows that

$$|a_n| = \frac{1}{n^2\pi} \left| \int_{-\pi}^{\pi} f''(x) \cos nx dx \right| < \frac{1}{n^2\pi} \int_{-\pi}^{\pi} M dx = \frac{2M}{n^2}$$

Similarly, $|b_n| < 2M/n^2$ for all n . Hence the absolute value of each term of the Fourier series corresponding to $f(x)$ is at most equal to the corresponding term of the series

$$|a_0| + 2M \left(1 + 1 + \frac{1}{2^2} + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{3^2} + \dots \right)$$

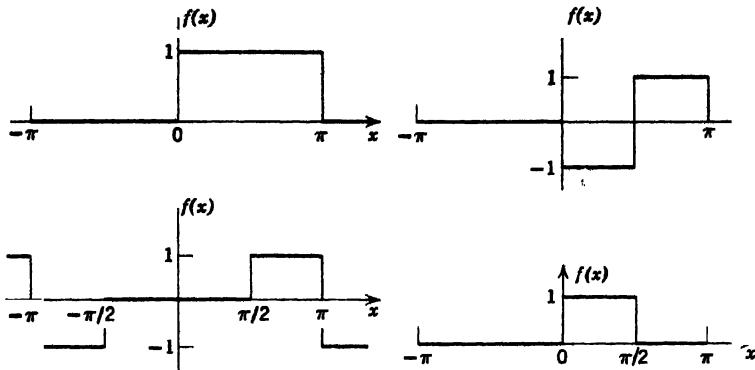
which is convergent. Hence that Fourier series converges and the proof is complete. (Readers already familiar with uniform convergence will see that, by the Weierstrass test in Sec. 14.4, the Fourier series converges uniformly, and our derivation of (6) by integrating term-by-term is then justified by Theorem 3 of that section.)

The proofs of convergence in the case of a piecewise continuous function $f(x)$ and of the last statement of Theorem 1 can be found in more advanced texts, for example in Ref. [E3].

Problems

Find the Fourier series of the function $f(x)$ which is assumed to have the period 2π and plot accurate graphs of the first three partial sums,⁵ where $f(x)$ equals

⁵ That is, $a_0 + \sum_{n=1}^N (a_n \cos nx + b_n \sin nx)$ for $N = 1, 2, 3$.



5. $f(x) = x \quad (-\pi < x < \pi)$

7. $f(x) = x^3 \quad (-\pi < x < \pi)$

9. $f(x) = \begin{cases} \frac{1}{2}\pi + x & \text{if } -\pi < x < 0 \\ \frac{1}{2}\pi - x & \text{if } 0 < x < \pi \end{cases}$

11. $f(x) = \begin{cases} 1 & \text{if } -\pi/2 < x < \pi/2 \\ -1 & \text{if } \pi/2 < x < 3\pi/2 \end{cases}$

13. $f(x) = \begin{cases} x & \text{if } -\pi/2 < x < \pi/2 \\ 0 & \text{if } \pi/2 < x < 3\pi/2 \end{cases}$

15. $f(x) = \begin{cases} 0 & \text{if } -\pi < x < 0 \\ x^2 & \text{if } 0 < x < \pi \end{cases}$

6. $f(x) = x^2 \quad (-\pi < x < \pi)$

8. $f(x) = x^4 \quad (-\pi < x < \pi)$

10. $f(x) = \begin{cases} -x & \text{if } -\pi < x < 0 \\ x & \text{if } 0 < x < \pi \end{cases}$

12. $f(x) = \begin{cases} 0 & \text{if } -\pi < x < 0 \\ x & \text{if } 0 < x < \pi \end{cases}$

14. $f(x) = \begin{cases} x & \text{if } -\pi/2 < x < \pi/2 \\ \pi - x & \text{if } \pi/2 < x < 3\pi/2 \end{cases}$

16. $f(x) = \begin{cases} x^2 & \text{if } -\pi/2 < x < \pi/2 \\ \pi^2/4 & \text{if } \pi/2 < x < 3\pi/2 \end{cases}$

17. If $f(x)$ has the Fourier coefficients a_n, b_n , show that $kf(x)$, where k is a constant, has the Fourier coefficients ka_n, kb_n .
18. Show that if $f(x)$ has the Fourier coefficients a_n, b_n and $g(x)$ has the Fourier coefficients a_n^*, b_n^* , then $f(x) + g(x)$ has the Fourier coefficients $a_n + a_n^*, b_n + b_n^*$.
19. Using Probs. 17 and 18 and the results of Probs. 1 and 2, find the Fourier series in Prob. 4.
20. Verify the last statement in Theorem 1 concerning the discontinuities for the function in Prob. 11.

8.3 Even and Odd Functions

In Example 1 of the preceding section we wasted time and energy by calculating the a_n , which then turned out to be zero. This was disappointing, and we may ask whether this result could have been obtained without performing the integration: In fact, this is possible. We shall see that the a_n are zero because $f(x)$ in Ex. 1 is an odd function.

We first remember that a function $y = g(x)$ is said to be even if

$$g(-x) = g(x)$$

for all x .

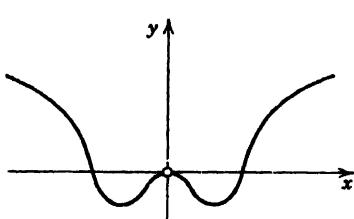


Fig. 219. Even function.

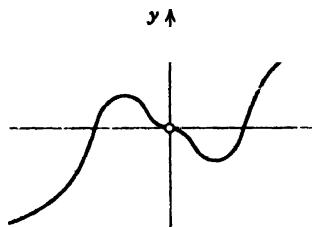


Fig. 220. Odd function.

The graph of such a function is symmetric with respect to the y -axis (Fig. 219). A function $h(x)$ is said to be **odd** if

$$h(-x) = -h(x) \quad \text{for all } x.$$

(Cf. Fig. 220.) The function $\cos nx$ is even, while $\sin nx$ is odd. (Cf. also Sec. 0.1.)

If $g(x)$ is an even function, then

$$(1) \quad \int_{-\pi}^{\pi} g(x) dx = 2 \int_0^{\pi} g(x) dx \quad (g \text{ even}).$$

If $h(x)$ is an odd function, then

$$(2) \quad \int_{-\pi}^{\pi} h(x) dx = 0 \quad (h \text{ odd}).$$

The formulas (1) and (2) are obvious from the graphs of g and h , and we leave the formal proofs to the student.

The product $q = gh$ of an even function g and an odd function h is odd, because

$$q(-x) = g(-x)h(-x) = g(x)[-h(x)] = -q(x).$$

Hence, if $f(x)$ is even, then $\int f \sin nx$ in (6c) of the last section is odd, and $b_n = 0$. Similarly, if $f(x)$ is odd, then $\int f \cos nx$ in (6b) is odd, and $a_n = 0$. From this and (1) we obtain

Theorem 1. The Fourier series of an even periodic function $f(x)$ having period 2π is a "Fourier cosine series"

$$(3) \quad f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos nx \quad (f \text{ even})$$

with coefficients

$$(4) \quad a_0 = \frac{1}{\pi} \int_0^{\pi} f(x) dx, \quad a_n = \frac{2}{\pi} \int_0^{\pi} f(x) \cos nx dx, \quad n = 1, 2,$$

The Fourier series of an odd periodic function $f(x)$ having period 2π is a "Fourier sine series"

$$(5) \quad f(x) = \sum_{n=1}^{\infty} b_n \sin nx \quad (f \text{ odd})$$

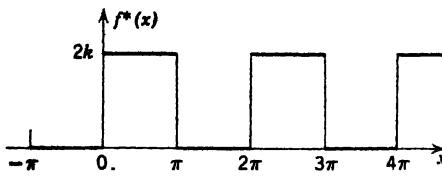


Fig. 221. Example 1.

with coefficients

$$(6) \quad b_n = \frac{2}{\pi} \int_0^\pi f(x) \sin nx \, dx.$$

For instance, $f(x)$ in Ex. 1, Sec. 8.2, is odd and, therefore, is represented by a Fourier sine series.

Further simplifications result from

Theorem 2. *The Fourier coefficients of a sum $f_1 + f_2$ are the sums of corresponding Fourier coefficients of f_1 and f_2 .*

Example 1 (Rectangular pulse). The function $f^*(x)$ in Fig. 221 is the sum of the function $f(x)$ in Ex. 1 of the last section and the constant k . Hence, from that example and Theorem 2 we conclude that

$$f^*(x) = k + \frac{4k}{\pi} \left(\sin x + \frac{1}{3} \sin 3x + \frac{1}{5} \sin 5x + \dots \right).$$

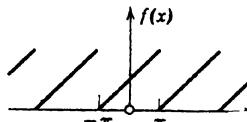
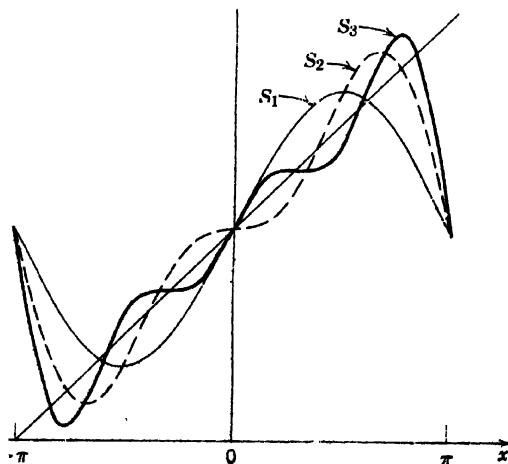
(a) The function $f(x)$.(b) Partial sums $S_n(x)$.

Fig. 222. Example 2.

Example 2. Find the Fourier series of the function (Fig. 222)

$$f(x) = x + \pi \quad \text{when } -\pi < x < \pi \quad \text{and} \quad f(x + 2\pi) = f(x).$$

We may write

$$f = f_1 + f_2 \quad \text{where } f_1 = x \quad \text{and} \quad f_2 = \pi.$$

The Fourier coefficients of f_2 are zero, except for the first one (the constant term), which is π . Hence, by Theorem 2, the Fourier coefficients a_n, b_n are those of f_1 , except for a_0 which is π . Since f_1 is odd, $a_n = 0$ for $n = 1, 2, \dots$, and

$$b_n = \frac{2}{\pi} \int_0^\pi f_1(x) \sin nx \, dx = \frac{2}{\pi} \int_0^\pi x \sin nx \, dx.$$

Integrating by parts we obtain

$$b_n = \frac{2}{\pi} \left[\frac{-x \cos nx}{n} \Big|_0^\pi + \frac{1}{n} \int_0^\pi \cos nx \, dx \right] = -\frac{2}{n} \cos n\pi.$$

Hence $b_1 = 2$, $b_2 = -2/2$, $b_3 = 2/3$, $b_4 = -2/4, \dots$, and

$$f(x) = \pi + 2 \left(\sin x - \frac{1}{2} \sin 2x + \frac{1}{3} \sin 3x - \dots \right).$$

Problems

1. Prove Theorem 2.

Are the following functions even, odd, or neither even nor odd?

2. $x + x^2$, $|x|$, e^x , e^{x^2} , $\sin^2 x$, $x \sin x$, $\ln x$, $x \cos x$
 3. $|x^3|$, $x \cos nx$, $x^2 \cos nx$, $\cosh x$, $\sinh x$, $\sin x + \cos x$, $x|x|$

Are the following functions $f(x)$ which are assumed to be periodic, of period 2π , even, odd, or neither even nor odd?

4. $f(x) = \begin{cases} 0 & \text{if } -\pi < x < 0 \\ x & \text{if } 0 < x < \pi \end{cases}$ 5. $f(x) = \begin{cases} -x^2 & \text{if } -\pi < x < 0 \\ x^2 & \text{if } 0 < x < \pi \end{cases}$
 6. $f(x) = \begin{cases} x & \text{if } -\pi/2 < x < \pi/2 \\ 0 & \text{if } \pi/2 < x < 3\pi/2 \end{cases}$ 7. $f(x) = \begin{cases} 0 & \text{if } -\pi < x < -\pi/2 \\ x & \text{if } -\pi/2 < x < \pi \end{cases}$
 8. $f(x) = |\sin x| \quad (-\pi < x < \pi)$ 9. $f(x) = e^{|x|} \quad (-\pi < x < \pi)$
 10. $f(x) = e^{-|x|} \quad (-\pi < x < \pi)$ 11. $f(x) = x|x| \quad (-\pi < x < \pi)$
 12. $f(x) = x \quad (0 < x < 2\pi)$ 13. $f(x) = x^2 \quad (0 < x < 2\pi)$

14. Find all functions which are both even and odd.

Prove

15. If $f(x)$ is odd, then $|f(x)|$ and $f^2(x)$ are even.
 16. If $f(x)$ is even, then $|f(x)|$, $f^2(x)$, and $f^3(x)$ are even.
 17. The sum of odd functions is odd. The product of two odd functions is even.
 18. The sum and the product of even functions are even.
 19. The product of an odd and an even function is odd.
 20. Show that if $g(x)$ is any function, defined for all x , then $p(x) = [g(x) + g(-x)]/2$ is even and $q(x) = [g(x) - g(-x)]/2$ is odd.

..

Represent the following functions as the sum of an even and an odd function.

21. e^x

22. $1/(1 - x)$

23. $x/(1 - x)$

24. $(1 + x)/(1 - x)$

Find the Fourier series of the following functions which are assumed to have the period 2π .

25. $f(x) = x \quad (-\pi < x < \pi)$

26. $f(x) = |x| \quad (-\pi < x < \pi)$

27. $f(x) = x^3 \quad (-\pi < x < \pi)$

28. $f(x) = |\sin x| \quad (-\pi < x < \pi)$

29. $f(x) = \begin{cases} x & \text{if } -\pi/2 < x < \pi/2 \\ \pi - x & \text{if } \pi/2 < x < 3\pi/2 \end{cases}$

30. $f(x) = \begin{cases} x & \text{if } 0 < x < \pi \\ \pi - x & \text{if } \pi < x < 2\pi \end{cases}$

31. $f(x) = \begin{cases} x - \pi & \text{if } 0 < x < \pi \\ -x & \text{if } \pi < x < 2\pi \end{cases}$

32. $f(x) = \begin{cases} -x^3 & \text{if } -\pi < x < 0 \\ x^3 & \text{if } 0 < x < \pi \end{cases}$

33. Show that $f(x) = x^3 \quad (-\pi < x \leq \pi)$, $f(x + 2\pi) = f(x)$ has the Fourier series

$$f(x) = \frac{\pi^2}{3} - 4 \left(\cos x - \frac{1}{4} \cos 2x + \frac{1}{9} \cos 3x - \dots \right).$$

34. Setting $x = \pi$ in Prob. 33, obtain the following famous result by Euler.

(7)
$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \dots = \frac{\pi^2}{6}.$$

35. Using Prob. 33, show that

(8)
$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} = 1 - \frac{1}{4} + \frac{1}{9} - \frac{1}{16} + \dots = \frac{\pi^2}{12}.$$

8.4 Functions Having Arbitrary Period

The transition from functions having period 2π to functions having any period T is quite simple, because it can be effected by a change of scale.

In fact, suppose that $f(t)$ has period T . Then we can introduce a new variable x such that $f(t)$, as a function of x , has period 2π . If we set

(1) (a) $t = \frac{T}{2\pi} x \quad \text{so that} \quad (b) \quad x = \frac{2\pi}{T} t,$

then $x = \pm\pi$ corresponds to $t = \pm T/2$, which means that f , as a function of x , has period 2π and, therefore, a Fourier series of the form

(2)
$$f(t) = f\left(\frac{T}{2\pi} x\right) = a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

whose coefficients are obtained from (6), Sec. 8.2, in the form

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f\left(\frac{T}{2\pi} x\right) dx, \quad a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f\left(\frac{T}{2\pi} x\right) \cos nx dx,$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f\left(\frac{T}{2\pi} x\right) \sin nx dx.$$

We could use these formulas directly, but the change to t simplifies calculation. Since

$$x = \frac{2\pi}{T} t, \quad \text{we have} \quad dx = \frac{2\pi}{T} dt,$$

and the interval of integration corresponds to the interval

$$-\frac{T}{2} \leq t \leq \frac{T}{2}.$$

Consequently, we obtain the Euler formulas

$$(3) \quad \begin{aligned} (a) \quad a_0 &= \frac{1}{T} \int_{-T/2}^{T/2} f(t) dt \\ (b) \quad a_n &= \frac{2}{T} \int_{-T/2}^{T/2} f(t) \cos \frac{2n\pi t}{T} dt \\ (c) \quad b_n &= \frac{2}{T} \int_{-T/2}^{T/2} f(t) \sin \frac{2n\pi t}{T} dt \quad n = 1, 2, \dots \end{aligned}$$

for the Fourier coefficients of $f(t)$. The Fourier series (2) with x expressed in terms of t becomes

$$(4) \quad f(t) = a_0 + \sum_{n=1}^{\infty} \left(a_n \cos \frac{2n\pi}{T} t + b_n \sin \frac{2n\pi}{T} t \right).$$

The interval of integration in (3) may be replaced by any interval of length T , for example, by the interval $0 \leq t \leq T$.

From Theorem 1 in Sec. 8.3 we now obtain

Theorem 1. *The Fourier series of an even function $f(t)$ having period T is a Fourier cosine series*

$$(5) \quad f(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos \frac{2n\pi}{T} t \quad (f \text{ even})$$

with coefficients

$$(6) \quad a_0 = \frac{2}{T} \int_0^{T/2} f(t) dt, \quad a_n = \frac{4}{T} \int_0^{T/2} f(t) \cos \frac{2n\pi}{T} t dt, \quad n = 1, 2, \dots$$

The Fourier series of an odd function $f(t)$ having period T is a Fourier sine series

$$(7) \quad f(t) = \sum_{n=1}^{\infty} b_n \sin \frac{2n\pi}{T} t \quad (f \text{ odd})$$

with coefficients

$$(8) \quad b_n = \frac{4}{T} \int_0^{T/2} f(t) \sin \frac{2n\pi}{T} t dt.$$

Example 1. Find the Fourier series of the function (cf. Fig. 223 on the next page)

$$f(t) = \begin{cases} 0 & \text{when } -2 < t < -1, \\ k & \text{when } -1 < t < 1, \\ 0 & \text{when } 1 < t < 2, \end{cases} \quad T = 4.$$

$$f(t) = \begin{cases} k & \text{when } -1 < t < 1, \\ 0 & \text{when } 1 < t < 2, \end{cases} \quad \dots$$

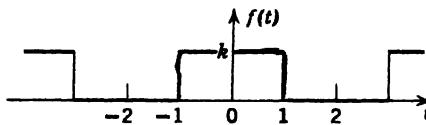


Fig. 223. Example 1.

Since f is even, $b_n = 0$. From (6),

$$a_0 = \frac{1}{2} \int_0^2 f(t) dt = \frac{1}{2} \int_0^1 k dt = \frac{k}{2},$$

$$a_n = \int_0^2 f(t) \cos \frac{n\pi}{2} t dt = \int_0^1 k \cos \frac{n\pi}{2} t dt = \frac{2k}{n\pi} \sin \frac{n\pi}{2}.$$

Thus $a_n = 0$ when n is even, $a_n = 2k/n\pi$ when $n = 1, 5, 9, \dots$, and $a_n = -2k/n\pi$ when $n = 3, 7, 11, \dots$. Hence

$$f(t) = \frac{k}{2} + \frac{2k}{\pi} \left(\cos \frac{\pi}{2} t - \frac{1}{3} \cos \frac{3\pi}{2} t + \frac{1}{5} \cos \frac{5\pi}{2} t - \dots \right).$$

Example 2 (Half-wave rectifier). A sinusoidal voltage $E \sin \omega t$ is passed through a half-wave rectifier which clips the negative portion of the wave (Fig. 224). Develop the resulting periodic function

$$u(t) = \begin{cases} 0 & \text{when } -T/2 < t < 0, \\ E \sin \omega t & \text{when } 0 < t < T/2, \end{cases} \quad T = \frac{2\pi}{\omega}$$

in a Fourier series.

Since $u = 0$ when $-T/2 < t < 0$, we obtain from (3a)

$$a_0 = \frac{\omega}{2\pi} \int_0^{\pi/\omega} E \sin \omega t dt = \frac{E}{\pi}$$

and from (3b), by using (11b) in Sec. 0.1, with $x = \omega t$ and $y = n\omega t$,

$$a_n = \frac{\omega}{\pi} \int_0^{\pi/\omega} E \sin \omega t \cos n\omega t dt = \frac{\omega E}{2\pi} \int_0^{\pi/\omega} [\sin(1+n)\omega t + \sin(1-n)\omega t] dt$$

When $n = 1$, the integral on the right is zero, and when $n = 2, 3, \dots$,

$$\begin{aligned} a_n &= \frac{\omega E}{2\pi} \left[-\frac{\cos(1+n)\omega t}{(1+n)\omega} - \frac{\cos(1-n)\omega t}{(1-n)\omega} \right]_0^{\pi/\omega} \\ &= \frac{E}{2\pi} \left(\frac{-\cos(1+n)\pi + 1}{1+n} + \frac{-\cos(1-n)\pi + 1}{1-n} \right). \end{aligned}$$

When n is odd, this is equal to zero, and for even n we obtain

$$a_n = \frac{E}{2\pi} \left(\frac{2}{1+n} + \frac{2}{1-n} \right) = -\frac{2E}{(n-1)(n+1)\pi} \quad (n = 2, 4, \dots).$$

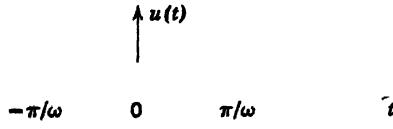


Fig. 224. Half-wave rectifier.

In a similar fashion we find from (3c) that $b_1 = E/2$ and $b_n = 0$ for $n = 2, 3, \dots$. Consequently,

$$u(t) = \frac{E}{\pi} + \frac{E}{2} \sin \omega t - \frac{2E}{\pi} \left(\frac{1}{1 \cdot 3} \cos 2\omega t + \frac{1}{3 \cdot 5} \cos 4\omega t + \dots \right).$$

Problems

1. Show that each term in (4) has the period T .
2. Show that in (3) the interval of integration may be replaced by any interval of length T .
3. Using (3), prove Theorem 1.

Find the Fourier series of the periodic function $f(t)$, of period T , and plot accurate graphs of $f(t)$ and the first three partial sums.

4. $f(t) = 0 \quad (-2 < t < 0), \quad f(t) = 1 \quad (0 < t < 2), \quad T = 4$
5. $f(t) = 1 \quad (-1 < t < 1), \quad f(t) = 0 \quad (1 < t < 3), \quad T = 4$
6. $f(t) = 1 \quad (-1 < t < 0), \quad f(t) = -1 \quad (0 < t < 1), \quad T = 2$
7. $f(t) = t \quad (-1 < t < 1), \quad T = 2$
8. $f(t) = |t| \quad (-2 < t < 2), \quad T = 4$
9. $f(t) = t \quad (0 < t < 1), \quad f(t) = 1 - t \quad (1 < t < 2), \quad T = 2$
10. $f(t) = 1/2 \quad (-1 < t < 0), \quad f(t) = -t \quad (0 < t < 1), \quad T = 2$
11. $f(t) = t \quad (-\pi/8 < t < \pi/8), \quad f(t) = (\pi/4) - t \quad (\pi/8 < t < 3\pi/8), \quad T = \pi/2$
12. $f(t) = (1/2) + t \quad (-1/2 < t < 0), \quad f(t) = (1/2) - t \quad (0 < t < 1/2), \quad T = 1$
13. $f(t) = t^2 \quad (-1 < t < 1), \quad T = 2$
14. $f(t) = \sin \pi t \quad (0 < t < 1), \quad T = 1$
15. $f(t) = 1 \quad (-2 < t < 0), \quad f(t) = e^{-t} \quad (0 < t < 2), \quad T = 4$

8.5 Half-Range Expansions

Let $f(t)$ have period $T = 2l$. If f is even, we obtain from Theorem 1 in Sec. 8.4 the Fourier cosine series

$$(1) \quad f(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi}{l} t \quad (f \text{ even})$$

with coefficients

$$(2) \quad a_0 = \frac{1}{l} \int_0^l f(t) dt, \quad a_n = \frac{2}{l} \int_0^l f(t) \cos \frac{n\pi}{l} t dt, \quad n = 1, 2, \dots$$

If f is odd, we obtain the Fourier sine series

$$(3) \quad f(t) = \sum_{n=1}^{\infty} b_n \sin \frac{n\pi}{l} t \quad (f \text{ odd})$$

with coefficients

$$(4) \quad b_n = \frac{2}{l} \int_0^l f(t) \sin \frac{n\pi}{l} t dt, \quad n = 1, 2, \dots$$

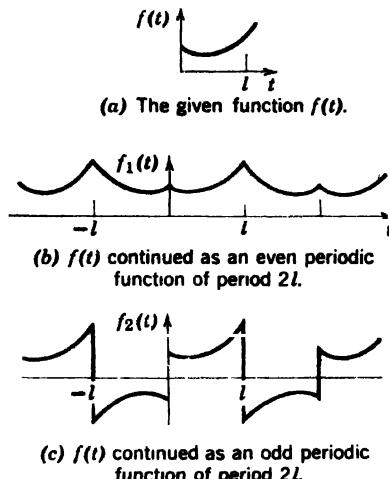


Fig. 225. Periodic continuations.

Now, (2) and (4) use only the values of $f(t)$ between $t = 0$ and $t = l$. Hence, for a function $f(t)$ given only over this interval, we can form the series (1) and (3). If $f(t)$ satisfies the conditions in Theorem 1, Sec. 8.2, both series will represent the given function in the interval $0 < t < l$. Outside this interval the series (1) will represent the even periodic extension or continuation of f having period $T = 2l$ (Fig. 225b) and (3) will represent the odd periodic continuation of f (Fig. 225c). The series (1) and (3) with coefficients given by (2) and (4) are called **half-range expansions** of the given function $f(t)$. They will have important applications in connection with partial differential equations (Secs. 9.3, 9.5).

Example 1. Find the half-range expansions of the function

$$f(t) = \begin{cases} \frac{2k}{l}t & \text{when } 0 < t < \frac{l}{2} \\ \frac{2k}{l}(l-t) & \text{when } \frac{l}{2} < t < l \end{cases}$$

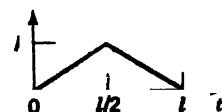


Fig. 226. The given function in Example 1.

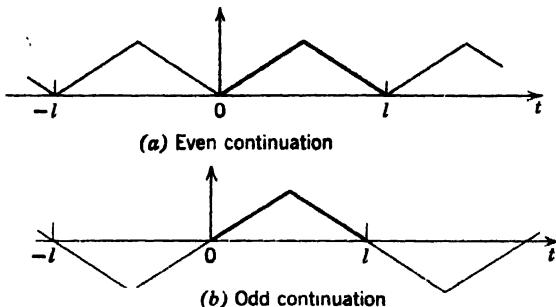
shown in Fig. 226. From (2),

$$a_0 = \frac{1}{l} \left[\frac{2k}{l} \int_0^{l/2} t dt + \frac{2k}{l} \int_{l/2}^l (l-t) dt \right] = \frac{k}{2},$$

$$a_n = \frac{2}{l} \left[\frac{2k}{l} \int_0^{l/2} t \cos \frac{n\pi}{l} t dt + \frac{2k}{l} \int_{l/2}^l (l-t) \cos \frac{n\pi}{l} t dt \right].$$

Now by integration by parts,

$$\begin{aligned} \int_0^{l/2} t \cos \frac{n\pi}{l} t dt &= \frac{lt}{n\pi} \sin \frac{n\pi}{l} t \Big|_0^{l/2} - \frac{l}{n\pi} \int_0^{l/2} \sin \frac{n\pi}{l} t dt \\ &= \frac{l^2}{2n\pi} \sin \frac{n\pi}{2} + \frac{l^2}{n^2\pi^2} \left(\cos \frac{n\pi}{2} - 1 \right). \end{aligned}$$

Fig. 227. Periodic continuations of $f(t)$ in Example 1.

Similarly,

$$\int_{l/2}^l (l-t) \cos \frac{n\pi}{l} t dt = -\frac{l^2}{2n\pi} \sin \frac{n\pi}{2} - \frac{l^2}{n^2\pi^2} \left(\cos n\pi - \cos \frac{n\pi}{2} \right).$$

By inserting these two results we obtain

$$a_n = \frac{4k}{n^2\pi^2} \left(2 \cos \frac{n\pi}{2} - \cos n\pi - 1 \right).$$

Thus,

$$a_2 = -16k/2^2\pi^2, \quad a_6 = -16k/6^2\pi^2, \quad a_{10} = -16k/10^2\pi^2, \dots$$

and $a_n = 0$ when $n \neq 2, 6, 10, 14, \dots$. Hence the first half-range expansion of $f(t)$ is

$$f(t) = \frac{k}{2} - \frac{16k}{\pi^2} \left(\frac{1}{2^2} \cos \frac{2\pi}{l} t + \frac{1}{6^2} \cos \frac{6\pi}{l} t + \dots \right).$$

This series represents the even periodic continuation of $f(t)$ shown in Fig. 227a.

Similarly, from (4),

$$(5) \quad b_n = \frac{8k}{n^2\pi^2} \sin \frac{n\pi}{2}$$

and the other half-range expansion is

$$f(t) = \frac{8k}{\pi^2} \left(\frac{1}{1^2} \sin \frac{\pi}{l} t - \frac{1}{3^2} \sin \frac{3\pi}{l} t + \frac{1}{5^2} \sin \frac{5\pi}{l} t - + \dots \right).$$

This series represents the odd periodic continuation of $f(t)$ shown in Fig. 227b.

Problems

Represent the following functions $f(x)$ by a Fourier cosine series and graph the corresponding periodic continuation of $f(x)$.

$$1. f(x) = 1 \quad (0 < x < l)$$

$$2. f(x) = x \quad (0 < x < l)$$

$$3. f(x) = \begin{cases} 1 & \text{if } 0 < x < l/2 \\ 0 & \text{if } l/2 < x < l \end{cases}$$

$$4. f(x) = \begin{cases} 0 & \text{if } 0 < x < l/2 \\ 1 & \text{if } l/2 < x < l \end{cases}$$

$$5. f(x) = 1 - \frac{x}{l} \quad (0 < x < l)$$

$$6. f(x) = x^2 \quad (0 < x < l)$$

$$7. f(x) = x^3 \quad (0 < x < l)$$

$$8. f(x) = e^x \quad (0 < x < l)$$

$$9. f(x) = \sin \frac{\pi}{l} x \quad (0 < x < l)$$

$$10. f(x) = \sin \frac{\pi x}{2l} \quad (0 < x < l)$$

Represent the following functions $f(x)$ by a Fourier sine series and graph the corresponding periodic continuation of $f(x)$.

11. $f(x) = 1 \quad (0 < x < l)$

12. $f(x) = x \quad (0 < x < l)$

13. $f(x) = \begin{cases} 1 & \text{if } 0 < x < l/2 \\ 0 & \text{if } l/2 < x < l \end{cases}$

14. $f(x) = \begin{cases} 1/2 & \text{if } 0 < x < l/2 \\ 3/2 & \text{if } l/2 < x < l \end{cases}$

15. $f(x) = x^2 \quad (0 < x < l)$

16. $f(x) = x^3 \quad (0 < x < l)$

17. (Complex form of the Fourier series, complex Fourier coefficients) Using the formula $e^{i\theta} = \cos \theta + i \sin \theta$ (cf. Sec. 2.4), show that

$$\cos nx = \frac{1}{2}(e^{inx} + e^{-inx}), \quad \sin nx = \frac{1}{2i}(e^{inx} - e^{-inx})$$

and the Fourier series

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

may be written in the form

$$(6) \quad f(x) = c_0 + \sum_{n=1}^{\infty} (c_n e^{inx} + k_n e^{-inx})$$

where $c_0 = a_0$, $c_n = (a_n - ib_n)/2$, $k_n = (a_n + ib_n)/2$, $n = 1, 2, \dots$. Using (6), Sec. 8.2, show that

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx, \quad k_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{inx} dx, \quad n = 1, 2, \dots$$

Introducing the notation $k_n = c_{-n}$, show that (6) may be written

$$(7) \quad f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}, \quad c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx, \quad n = 0, \pm 1, \pm 2, \dots$$

This is the so-called *complex form* of the Fourier series, and the c_n are called the *complex Fourier coefficients* of $f(x)$.

18. Using (7), show that the complex form of the Fourier series of the function

$$f(x) = e^x \quad \text{when} \quad -\pi < x < \pi \quad \text{and} \quad f(x + 2\pi) = f(x)$$

$$f(x) = \frac{\sinh \pi}{\pi} \sum_{n=-\infty}^{\infty} (-1)^n \frac{1 + i n}{1 + n^2} e^{inx}.$$

19. Obtain from the series in Prob. 18 the real Fourier series

$$f(x) = K \left[\frac{1}{2} - \frac{1}{1 + 1^2} (\cos x - \sin x) + \frac{1}{1 + 2^2} (\cos 2x - 2 \sin 2x) - + \dots \right].$$

where $K = (2 \sinh \pi)/\pi$.

20. Show that the complex Fourier coefficients of an odd function are purely imaginary and those of an even function are real.

8.6 Determination of Fourier Coefficients without Integration

In the previous sections relatively complicated and lengthy integrations led to relatively simple expressions for the Fourier coefficients a_n and b_n . This raises the question whether there might be a simpler way of obtaining Fourier coefficients. There is, and we want to show that the Fourier coefficients of a periodic function which is represented by polynomials can be obtained in terms of the jumps of the function and its derivatives. Of course, this is a big advantage, and the corresponding formulas are of great practical importance, because by applying them integrations are avoided (except for a_0 , which must be determined as before).

By a *jump* j of a function $g(x)$ at a point x_0 we mean the difference between the right-hand and left-hand limits (Sec. 8.2) of $g(x)$ at x_0 ; that is (Fig. 228)

$$(1) \quad j = g(x_0 + 0) - g(x_0 - 0).$$

It follows that an upward jump is positive and a downward jump is negative.

Let $f(x)$ be a function which has period 2π and is represented by polynomials p_1, \dots, p_m in the interval $-\pi < x < \pi$, say (Fig. 229)

$$(2) \quad f(x) = \begin{cases} p_1(x) & \text{when } x_0 < x < x_1, \quad (x_0 = -\pi) \\ p_2(x) & \text{when } x_1 < x < x_2, \\ & \vdots \\ & \vdots \\ p_m(x) & \text{when } x_{m-1} < x < x_m (= \pi). \end{cases}$$

Then f may have jumps at x_0, x_1, \dots, x_m , and the same is true for the derivatives f', f'', \dots . We choose the following notation.

$$(3) \quad \begin{aligned} j_s &= \text{jump of } f \text{ at } x_s, \\ j'_s &= \text{jump of } f' \text{ at } x_s, \quad (s = 1, 2, \dots, m) \\ j''_s &= \text{jump of } f'' \text{ at } x_s, \quad \text{etc.} \end{aligned}$$

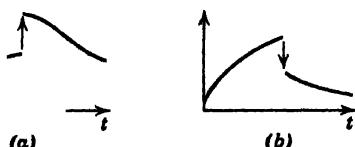


Fig. 228. Jump of a function. (a) Positive jump. (b) Negative jump.

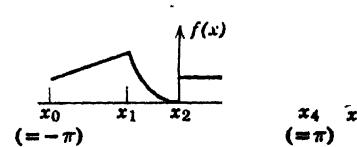


Fig. 229. Example of a representation of the form (2) (with $m = 4$).
• •

Of course, if f is continuous at x_s , then $j_s = 0$, and for the derivatives the situation is similar, so that some of the numbers j_s, j'_s, \dots in (3) may be zero.

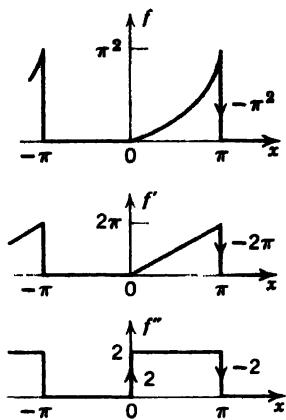


Fig. 230. $f(x)$ and derivatives in Example 1.

Example 1. Let

$$f(x) = \begin{cases} 0 & \text{when } -\pi < x < 0, \\ x^2 & \text{when } 0 < x < \pi. \end{cases}$$

We plot graphs of f and its derivatives (Fig. 230),

$$f' = \begin{cases} 0 & \text{when } -\pi < x < 0, \\ 2x & \text{when } 0 < x < \pi. \end{cases} \quad f'' = \begin{cases} 0 & \text{when } -\pi < x < 0, \\ 2 & \text{when } 0 < x < \pi. \end{cases} \quad f''' = 0.$$

We see that the jumps are

	Jump at $x_1 = 0$	Jump at $x_2 = \pi$
f	$j_1 = 0$	$j_2 = -\pi^2$
f'	$j'_1 = 0$	$j'_2 = -2\pi$
f''	$j''_1 = 2$	$j''_2 = -2$

Note that the jumps at $x = -\pi$ are not listed, because they are taken into account at $x = \pi$, the other end of the interval of periodicity.

To derive the desired formula for the Fourier coefficients a_1, a_2, \dots of f , given by (2), we start from the Euler formula (6b) in Sec. 8.2:

$$(4) \quad \pi a_n = \int_{-\pi}^{\pi} f \cos nx dx.$$

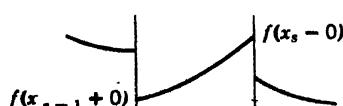
Since f is represented by (2), we write the integral as the sum of m integrals:

$$(5) \quad \pi a_n = \int_{x_0}^{x_1} + \int_{x_1}^{x_2} + \cdots + \int_{x_{m-1}}^{x_m} = \sum_{s=1}^m \int_{x_{s-1}}^{x_s} f \cos nx dx,$$

where $x_0 = -\pi$ and $x_m = \pi$. Integration by parts yields

$$(6) \quad \int_{x_{s-1}}^{x_s} f \cos nx dx = \frac{f}{n} \sin nx \Big|_{x_{s-1}}^{x_s} - \frac{1}{n} \int_{x_{s-1}}^{x_s} f' \sin nx dx.$$

Now comes an important point: the evaluation of the first expression on the right. $f(x)$ may be discontinuous at x_s (Fig. 231), and we have to take the left-hand limit $f(x_s - 0)$ of f at x_s . Similarly, at x_{s-1} we have to take the right-hand limit $f(x_{s-1} + 0)$. Hence the first expression on the right-hand side of (6) equals



$$\frac{1}{n} [f(x_s - 0) \sin nx_s - f(x_{s-1} + 0) \sin nx_{s-1}].$$

Fig. 231. The formula (6).

Consequently, by inserting (6) into (5) and using the short notations $S_0 = \sin nx_0$, $S_1 = \sin nx_1$, etc., we obtain

$$(7) \quad \begin{aligned} \pi a_n &= \frac{1}{n} [f(x_1 - 0)S_1 - f(x_0 + 0)S_0 + f(x_2 - 0)S_2 - f(x_1 + 0)S_1 \\ &\quad + \cdots + f(x_m - 0)S_m - f(x_{m-1} + 0)S_{m-1}] \\ &\quad - \frac{1}{n} \sum_{s=1}^m \int_{x_{s-1}}^{x_s} f' \sin nx \, dx. \end{aligned}$$

If we collect terms with the same S , the expression in brackets becomes

$$(8) \quad \begin{aligned} &-f(x_0 + 0)S_0 + [f(x_1 - 0) - f(x_1 + 0)]S_1 \\ &\quad + [f(x_2 - 0) - f(x_2 + 0)]S_2 + \cdots + f(x_m - 0)S_m. \end{aligned}$$

The expressions in brackets in (8) are the jumps of f , multiplied by -1 . Furthermore, because of periodicity, $S_0 = S_m$ and $f(x_0) = f(x_m)$, so that we may combine the first and the last term in (8). Thus (8) is equal to

$$-j_1 S_1 - j_2 S_2 - \cdots - j_m S_m,$$

and from (7) we therefore have the intermediate result

$$(9) \quad \pi a_n = -\frac{1}{n} \sum_{s=1}^m j_s \sin nx_s - \frac{1}{n} \sum_{s=1}^m \int_{x_{s-1}}^{x_s} f' \sin nx \, dx.$$

By applying the same procedure to the integrals on the right-hand side of (9) we find

$$(10) \quad \sum_{s=1}^m \int_{x_{s-1}}^{x_s} f' \sin nx \, dx = \frac{1}{n} \sum_{s=1}^m j_s' \cos nx_s + \frac{1}{n} \sum_{s=1}^m \int_{x_{s-1}}^{x_s} f'' \cos nx \, dx.$$

Continuing in this fashion, we obtain integrals involving higher and higher derivatives of f . Since f is represented by polynomials and the $(r+1)$ th derivative of a polynomial of degree r is identically zero, we shall reach the point where no integrals are left, and this will happen after finitely many steps. By inserting (10) and the analogous formulas obtained in the further steps into (9) we obtain the desired formula

$$(11a) \quad \begin{aligned} a_n &= \frac{1}{n\pi} \left[-\sum_{s=1}^m j_s \sin nx_s - \frac{1}{n} \sum_{s=1}^m j_s' \cos nx_s \right. \\ &\quad \left. + \frac{1}{n^2} \sum_{s=1}^m j_s'' \sin nx_s + \frac{1}{n^3} \sum_{s=1}^m j_s''' \cos nx_s - \cdots + \cdots \right]. \end{aligned}$$

In precisely the same fashion we obtain from (6c) in Sec. 8.2

$$(11b) \quad \begin{aligned} b_n &= \frac{1}{n\pi} \left[\sum_{s=1}^m j_s \cos nx_s - \frac{1}{n} \sum_{s=1}^m j_s' \sin nx_s \right. \\ &\quad \left. - \frac{1}{n^2} \sum_{s=1}^m j_s'' \cos nx_s + \frac{1}{n^3} \sum_{s=1}^m j_s''' \sin nx_s + \cdots + \cdots \right]. \end{aligned}$$

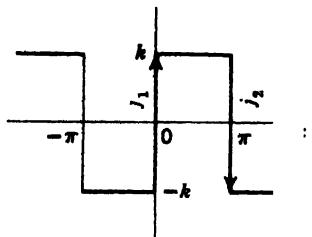


Fig. 232. Example 2.

To avoid errors it is practical to graph f' and its derivatives and list the jumps in a tab as in Ex. 1. The coefficient a_0 must be obtained by integration as before.

Example 2. Find the Fourier coefficients of the function (cf. Fig. 232)

$$f(x) = \begin{cases} -k & \text{when } -\pi < x < 0 \\ k & \text{when } 0 < x < \pi \end{cases}$$

We see that $f' \equiv 0$ and the jumps of f are

	Jump at $x_1 = 0$	Jump at $x_2 = \pi$
f	$j_1 = 2k$	$j_2 = -2k$

f is odd. Hence $a_n = 0$, and from (11b),

$$b_n = \frac{1}{n\pi} [j_1 \cos nx_1 + j_2 \cos nx_2] = \frac{1}{n\pi} [2k \cos 0 - 2k \cos n\pi]$$

$$= \frac{2k}{n\pi} (1 - \cos n\pi) = \begin{cases} 4k/n\pi & \text{for odd } n \\ 0 & \text{for even } n \end{cases} \quad (\text{cf. Ex. 1, Sec. 8.2}).$$

Example 3. Find the Fourier series of the function in Ex. 1 of the current section. By integration,

$$a_0 = \frac{1}{2\pi} \int_0^\pi x^2 dx = \frac{\pi^2}{6}.$$

From (11a) we obtain

$$a_n = \frac{1}{n\pi} \left[\pi^2 \sin n\pi + \frac{2\pi}{n} \cos n\pi + \frac{1}{n^2} (2 \sin 0 - 2 \sin n\pi) \right] = \frac{2}{n^2}$$

Hence $a_1 = -2/1^2$, $a_2 = 2/2^2$, $a_3 = -2/3^2$, \dots . From (11b) it follows that

$$b_n = \frac{1}{n\pi} \left[-\pi^2 \cos n\pi + \frac{2\pi}{n} \sin n\pi - \frac{1}{n^2} (2 \cos 0 - 2 \cos n\pi) \right]$$

$$= -\frac{\pi}{n} \cos n\pi + \frac{2}{n^3\pi} (\cos n\pi - 1).$$

Hence,

$$b_1 = \pi - \frac{4}{\pi}, \quad b_2 = -\frac{\pi}{2}, \quad b_3 = \frac{\pi}{3} - \frac{4}{3^3\pi}, \quad b_4 = -\frac{\pi}{4}, \dots$$

and the Fourier series is

$$f(x) = \frac{\pi^2}{6} - 2 \cos x + \left(\pi - \frac{4}{\pi} \right) \sin x + \frac{1}{2} \cos 2x - \frac{\pi}{2} \sin 2x + \dots$$

Problems

1. Derive (11b) from (6c), Sec. 8.2.

Using (11), find the Fourier series of the functions $f(x)$ in

2. Probs. 1–4, Sec. 8.2

3. Probs. 9–12, Sec. 8.2

4. Probs. 5–8, Sec. 8.2

5. Probs. 13–16, Sec. 8.2

6. Show that in the case of a function $f(t)$ having period T the formulas corresponding to (11) are

$$(12a) \quad \begin{aligned} 1 & \sum_{s=1}^m j_s \sin \frac{2n\pi}{T} t_s - \frac{T}{2n\pi} \sum_{s=1}^m j_s' \cos \frac{2n\pi}{T} t_s + \left(\frac{T}{2n\pi} \right)^2 \sum_{s=1}^m j_s'' \sin \frac{2n\pi}{T} t_s \\ & + \left(\frac{T}{2n\pi} \right)^3 \sum_{s=1}^m j_s''' \cos \frac{2n\pi}{T} t_s - - + + \cdots \end{aligned}$$

$$(12b) \quad \begin{aligned} b_n = \frac{1}{n\pi} \left[& \sum_{s=1}^m j_s \cos \frac{2n\pi}{T} t_s - \frac{T}{2n\pi} \sum_{s=1}^m j_s' \sin \frac{2n\pi}{T} t_s - \left(\frac{T}{2n\pi} \right)^2 \sum_{s=1}^m j_s'' \cos \frac{2n\pi}{T} t_s \right. \\ & \left. + \left(\frac{T}{2n\pi} \right)^3 \sum_{s=1}^m j_s''' \sin \frac{2n\pi}{T} t_s - - + + \cdots \right]. \end{aligned}$$

Using (12), find the Fourier series of the functions $f(t)$ in:

7. Probs. 4–6, Sec. 8.4

8. Probs. 7, 8, Sec. 8.4

9. Probs. 9–11, Sec. 8.4

10. Probs. 12, 13, Sec. 8.4

Using (12), find the Fourier cosine series of the functions $f(x)$ in

11. Probs. 1–4, Sec. 8.5

12. Probs. 5–7, Sec. 8.5

Using (12), find the Fourier sine series of the functions $f(x)$ in

13. Probs. 11–13, Sec. 8.5

14. Probs. 14–16, Sec. 8.5

15. Can (11) be applied to find the Fourier coefficients of the function $f(x) = e^x$ ($0 < x < 2\pi$), $f(x + 2\pi) = f(x)$?

8.7 Forced Oscillations

Fourier series have important applications in connection with differential equations. Let us consider an important practical problem involving an ordinary differential equation. (Partial differential equations will be considered in the next chapter.)

From Sec. 2.13 we know that forced oscillations of a body of mass m on a spring (modulus k) are governed by the equation

$$(1) \quad m\ddot{y} + c\dot{y} + ky = r(t).$$

If the external force $r(t)$ is a sine or cosine function and the damping constant c is not zero, the steady-state solution represents a harmonic oscillation having the frequency of the external force.

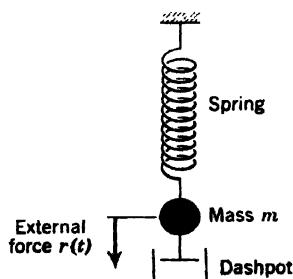


Fig. 233. Vibrating system under consideration.

We shall now see that if $r(t)$ is not a pure sine or cosine function but any other periodic function, then the steady-state solution will represent a superposition of harmonic oscillations having the frequency of $r(t)$ and multiples of this frequency. If the frequency of one of these oscillations is close to the resonant frequency of the vibrating system (cf. Sec. 2.13), then that oscillation may be the dominant part of the response of the system to the external force. Of course, this is quite surprising to an observer not familiar with the corresponding mathematical theory which is highly important in the study of vibrating systems and resonance. Let us illustrate the situation by an example.

Example 1. In (1), let $m = 1$ (gm), $c = 0.02$ (gm/sec), and $k = 25$ (gm/sec 2), so that (1) becomes

$$(2) \quad \ddot{y} + 0.02\dot{y} + 25y = r(t)$$

where $r(t)$ is measured in gm · cm/sec 2 . Let (Fig. 234)

$$r(t) = \begin{cases} t + \frac{\pi}{2} & \text{when } -\pi < t < 0, \\ -t + \frac{\pi}{2} & \text{when } 0 < t < \pi, \end{cases} \quad r(t + 2\pi) = r(t).$$

Find the steady-state solution $y(t)$.

We represent $r(t)$ by a Fourier series, finding

$$(3) \quad r(t) = \frac{4}{\pi} \left(\cos t + \frac{1}{3^2} \cos 3t + \frac{1}{5^2} \cos 5t + \dots \right).$$

Then we consider the differential equation

$$(4) \quad \ddot{y} + 0.02\dot{y} + 25y = \frac{4}{n^2\pi} \cos nt \quad (n = 1, 3, \dots)$$

whose right-hand side is a single term of the series (3).

From Sec. 2.13 we know that the steady-state solution $y_n(t)$ of (4) is of the form

$$(5) \quad y_n = A_n \cos nt + B_n \sin nt,$$

and by substituting this into (4) we find that

$$(6) \quad A_n = \frac{4(25 - n^2)}{n^2\pi D}, \quad B_n = \frac{0.08}{n\pi D} \quad \text{where} \quad D = (25 - n^2)^2 + (0.02n)^2.$$

Since (2) is linear, we may expect that the steady-state solution is

$$(7) \quad y = y_1 + y_3 + y_5 + \dots$$

where y_n is given by (5) and (6). In fact, this follows readily by substituting (7) into (2) and using the Fourier series of $r(t)$, provided that termwise differentiation of (7) is permissible. (Readers already familiar with the notion of uniform convergence [Sec. 14.4] may prove that (7) may be differentiated term by term.)

From (6) we find that the amplitude of (5) is

$$C_n = \sqrt{A_n^2 + B_n^2} = \frac{4}{n^2\pi\sqrt{D}}.$$

For $n = 5$ the quantity D is very small, the denominator of C_5 is small, and C_5 is so large that y_5 is the dominating term in (7). This implies that the steady-state motion is almost a harmonic oscillation whose frequency equals five times that of the exciting force (Fig. 235).

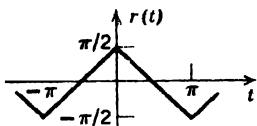


Fig. 234. Force in Example 1.

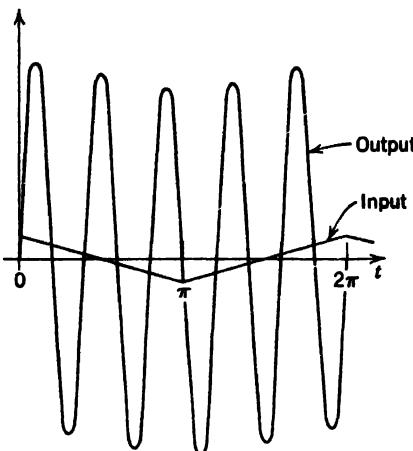


Fig. 235. Input and steady state output in Example 1.

Table 18. Amplitudes in Example 1.

n	C_n
1	0.053
3	0.0088
5	0.51
7	0.0011
9	0.00028
11	0.00011

Problems

Find the general solution of the differential equation $y'' + \omega^2y = r(t)$ where:

1. $r(t) = \sin t$, $\omega = 0.5, 0.7, 0.9, 1.1, 1.5, 2.0, 10.0$
2. $r(t) = \cos \alpha t + \cos \beta t$ ($\omega^2 \neq \alpha^2, \beta^2$)
3. $r(t) = \sin t + \frac{1}{8} \sin 3t + \frac{1}{24} \sin 5t$, $\omega = 0.5, 0.9, 1.1, 2, 2.9, 3.1, 4, 4.9, 5.1, 6, 8$
4. $r(t) = \begin{cases} t & \text{if } -\pi/2 < t < \pi/2 \\ \pi - t & \text{if } \pi/2 < t < 3\pi/2 \end{cases}$ and $r(t + 2\pi) = r(t)$, $|\omega| \neq 1, 3, 5, \dots$
5. $r(t) = \frac{t}{4}$ when $-\pi < t < \pi$ and $r(t + 2\pi) = r(t)$, $|\omega| \neq 0, 1, 2, \dots$
6. $r(t) = \sum_{n=1}^N a_n \cos nt$, $|\omega| \neq 1, 2, \dots, N$
7. $r(t) = \frac{\pi}{4} |\sin t|$ when $-\pi < t < \pi$ and $r(t + 2\pi) = r(t)$, $|\omega| \neq 0, 2, 4, \dots$
8. $r(t) = \begin{cases} t + \pi & \text{if } -\pi < t < 0 \\ -t + \pi & \text{if } 0 < t < \pi \end{cases}$ and $r(t + 2\pi) = r(t)$, $|\omega| \neq 0, 1, 3, \dots$

Find the steady-state oscillation corresponding to $y + cy + y = r(t)$ where $c > 0$ and

9. $r(t) = K \sin t$

10. $r(t) = \sin 3t$

11. $r(t) = a_n \cos nt$

12. $r(t) = \sum_{n=1}^N b_n \sin nt$

13. $r(t) = \frac{t}{12}(\pi^2 - t^2)$ when $-\pi < t < \pi$ and $r(t + 2\pi) = r(t)$

14. Find the steady-state current $I(t)$ in the RLC -circuit in Fig. 236 where $R = 100$ ohm, $L = 10$ henrys, $C = 10^{-2}$ farads, $E(t) = 100t(\pi^2 - t^2)$ volts when $-\pi < t < \pi$ and $E(t + 2\pi) = E(t)$. Proceed as follows. Develop $E(t)$ in a Fourier series. $I(t)$ will appear in the form of a trigonometric series. Find the general formulas for the coefficients of this series. Compute numerical values of the first few coefficients with a slide rule. Graph the sum of the first few terms of that series.

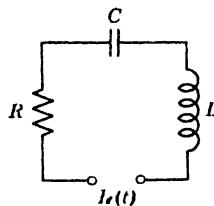


Fig. 236. RLC -circuit

15. Same task as in Prob. 14 when

$$E(t) = \begin{cases} 100(\pi t + t^2) & \text{if } -\pi < t < 0 \\ 100(\pi t - t^2) & \text{if } 0 < t < \pi \end{cases} \quad \text{and} \quad E(t + 2\pi) = E(t)$$

the other data being as before

8.8 Numerical Methods for Determining Fourier Coefficients. Square Error

The Fourier coefficients of a function $f(x)$ having period 2π are given by the Euler formulas (6), Sec. 8.2, and the occurring integrals may be evaluated by numerical integration, for example, by the rectangular rule (Sec. 0.8).

For this purpose we subdivide the interval of integration $-\pi \leq x \leq \pi$ into q equal parts of length $\Delta x = 2\pi/q$, whose common end points will be denoted by

$$x_0 (= -\pi), \quad x_1, \quad x_2, \quad \dots, \quad x_q (= \pi)$$

While in (2), Sec. 0.8, the function values $f(x_m^*)$ at the midpoints x_m^* of the intervals of subdivision are chosen, we shall now simply choose the function value $f(x_m)$ at the right end point of each interval. Since the integrals in (6),

Sec. 8.2, are multiplied by $1/2\pi$ and $1/\pi$, and $\Delta x/2\pi = 1/q$, $\Delta x/\pi = 2/q$, we thus obtain from (6), Sec. 8.2, the formulas

$$(1) \quad \begin{aligned} a_0 &= \frac{1}{q} [f(x_1) + \dots + f(x_q)] \\ a_n &= \frac{2}{q} [f(x_1) \cos nx_1 + \dots + f(x_q) \cos nx_q] \\ b_n &= \frac{2}{q} [f(x_1) \sin nx_1 + \dots + f(x_q) \sin nx_q] \end{aligned} \quad (n = 1, 2, \dots)$$

Since the values of sine and cosine are repeated in the four quadrants, it is practical to choose for q a number divisible by 4. A good choice is $q = 24$ because then many terms of the sums in (1) are zero. Let us illustrate the practical procedure by a simple example.

Example 1. Taking $q = 24$ in (1), compute a_0, \dots, a_{12} in the case of the function (cf. Prob. 33 in Sec. 8.3)

$$f(x) = x^2 \quad \text{when} \quad -\pi < x < \pi \quad \text{and} \quad f(x + 2\pi) = f(x)$$

Here $f(x_m) = x_m^2$. From (1) and Table 19,

$$a_0 = \frac{1}{24} (x_1^2 + \dots + x_{24}^2) = \frac{1}{24} (S_1 + 2S_2) = 3301 \quad \left(\text{exact } \frac{\pi^2}{3} \approx 3290 \right)$$

The values of the cosine required in the computation of the further coefficients are contained in Table 20 on the next page, and (1) yields

$$\begin{aligned} a_1 &= \frac{1}{12} [x_1^2 \cos(-165^\circ) + x_2^2 \cos(-150^\circ) + \dots] \\ &= \frac{1}{12} [x_1^2(-C_1) + x_2^2(-C_2) + \dots + x_8^2 C_6 + x_7^2 C_5 + \dots] \end{aligned}$$

Table 19. Values of $f(x)$ in Example 1.

m	x_m		x_m^2
	degrees	radians	
24	± 180	$\pm \pi$	9 869 60
1 23	± 165	$\pm 11\pi/12$	8 293 21
2 22	± 150	$\pm 5\pi/6$	6 853 89
3 21	± 135	$\pm 3\pi/4$	5 551 65
4 20	± 120	$\pm 2\pi/3$	4 386 49
5 19	± 105	$\pm 7\pi/12$	3 358 41
6 18	± 90	$\pm \pi/2$	2 467 40
7 17	± 75	$\pm 5\pi/12$	1 713 47
8 16	± 60	$\pm \pi/3$	1 096 62
9 15	± 45	$\pm \pi/4$	0 616 85
10 14	± 30	$\pm \pi/6$	0 274 16
11 13	± 15	$\pm \pi/12$	0 068 54
12	0	0	0

$$\text{Sums } S_x = 9869.60 \quad S_x^2 = 34680.69$$

Table 20. Values of $\cos x$ in Example 1.

x	$\cos x$	Notation
0°	1	C_0
15°	0.965 925 83	C_1
30°	0.866 025 40	C_2
45°	0.707 106 78	C_3
60°	0.5	C_4
75	0.258 819 05	C_5
90°	0	C_6

Collecting terms containing like C , omitting terms which are zero, and using $x_m^2 = x_{24-m}^2$, we find

$$a_1 = -\frac{1}{6} [x_0^2 C_0/2 + (x_1^2 - x_{11}^2) C_1 + (x_2^2 - x_{10}^2) C_2 + \cdots + (x_5^2 - x_7^2) C_5].$$

Using the function values in Table 20, we obtain the numerical result

$$= -4.023 \quad (\text{exact: } -4; \text{ cf. Prob. 33 in Sec. 8.3}).$$

The other coefficients a_2, \dots, a_{12} may be computed from (1) in a similar fashion. The result is shown in Table 21.

Table 21. Fourier Coefficients in Example 1.

n	a_n	
	by (1)	exact
0	3.301	3.290
1	-4.023	-4.000
2	1.023	1.000
3	-0.468	-0.444
4	0.274	0.250
5	-0.185	-0.160
6	0.137	0.111
7	-0.109	-0.082
8	0.091	0.063
9	-0.080	-0.049
10	0.073	0.040
11	-0.070	-0.033
12	0.069	0.028

We want to mention that there is another less naive approach to the formulas (1) as follows. Let $f(x)$ be a given function that has period 2π and can be represented by a Fourier series. Then the N th partial sum is an approximation for $f(x)$:

$$(2) \quad f(x) \approx a_0 + \sum_{n=1}^N (a_n \cos nx + b_n \sin nx).$$

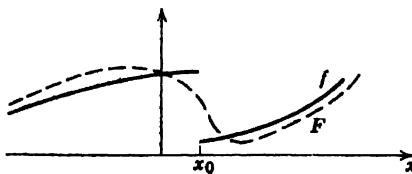


Fig. 237. Error of approximation.

We ask whether we can find a trigonometric polynomial

$$(3) \quad F(x) = \alpha_0 + \sum_{n=1}^N (\alpha_n \cos nx + \beta_n \sin nx)$$

which, for the same fixed N , is a "better" approximation for $f(x)$, that is, an approximation whose error is smaller than that of (2).

Of course, we must first define what we mean by an error of such an approximation. We want to choose a definition which measures the goodness of agreement between f and F on the whole interval $-\pi \leq x \leq \pi$. Obviously, the maximum of $|f - F|$ is not suitable for that purpose: in Fig. 237, the function F is a good approximation of f , but $|f - F|$ is large near x_0 . We choose for our consideration

$$(4) \quad E = \int_{-\pi}^{\pi} (f - F)^2 dx.$$

This is the so-called total square error of F relative to f over the interval $-\pi \leq x \leq \pi$. Clearly, $E \geq 0$.

N being fixed, we want to determine the coefficients in (3) such that E is minimum. We can write (4) in the form

$$(5) \quad E = \int_{-\pi}^{\pi} f^2 dx - 2 \int_{-\pi}^{\pi} fF dx + \int_{-\pi}^{\pi} F^2 dx.$$

By inserting (3) into the last integral and evaluating the occurring integrals as in Sec. 8.2 we readily obtain

$$\int_{-\pi}^{\pi} F^2 dx = \pi(2\alpha_0^2 + \alpha_1^2 + \cdots + \alpha_N^2 + \beta_1^2 + \cdots + \beta_N^2).$$

By inserting (3) into the second integral in (5) we see that the occurring integrals are those in the Euler formulas (6), Sec. 8.2, and

$$\int_{-\pi}^{\pi} fF dx = \pi(2\alpha_0 a_0 + \alpha_1 a_1 + \cdots + \alpha_N a_N + \beta_1 b_1 + \cdots + \beta_N b_N).$$

With these expressions (5) becomes

$$(6) \quad E = \int_{-\pi}^{\pi} f^2 dx - 2\pi \left[2\alpha_0 a_0 + \sum_{n=1}^N (\alpha_n a_n + \beta_n b_n) \right] + \pi \left[2\alpha_0^2 + \sum_{n=1}^N (\alpha_n^2 + \beta_n^2) \right].$$

If we take $\alpha_n = a_n$ and $\beta_n = b_n$ in (3), then from (6) we see that the square error corresponding to this particular choice of the coefficients of F is

$$(7) \quad E^* = \int_{-\pi}^{\pi} f^2 dx - \pi \left[2a_0^2 + \sum_{n=1}^N (a_n^2 + b_n^2) \right].$$

By subtracting (7) from (6) we obtain

$$E - E^* = \pi \left\{ 2(\alpha_0 - a_0)^2 + \sum_{n=1}^N [(\alpha_n - a_n)^2 + (\beta_n - b_n)^2] \right\}.$$

The terms on the right are squares of real numbers and therefore non-negative. Hence

$$E - E^* \geq 0 \quad \text{or} \quad E \geq E^*,$$

and $E = E^*$ if, and only if, $\alpha_0 = a_0, \dots, \beta_N = b_N$. This proves

Theorem 1. *The total square error of F [cf. (3), N fixed] relative to f on the interval $-\pi \leq x \leq \pi$ is minimum if and only if the coefficients of F in (3) are the corresponding Fourier coefficients of f . This minimum value is given by (7).*

The reader may show that the result $\alpha_0 = a_0, \dots, \beta_N = b_N$ may also be obtained from the conditions (10) (below) which are necessary (but not sufficient) for the minimum.

Let us look at (7). We see that E^* cannot increase as N increases, but may decrease. Hence, with increasing N , the partial sums of the Fourier series of f yield better and better approximations for $f(x)$, considered from the viewpoint of the square error.

Since $E^* \geq 0$ and (7) holds for any N , we obtain the important Bessel inequality⁶

$$(8) \quad 2a_0^2 + \sum_{n=1}^{\infty} (a_n^2 + b_n^2) \leq \frac{1}{\pi} \int_{-\pi}^{\pi} f^2(x) dx$$

for the Fourier coefficients of any function f for which the integral on the right exists.

In connection with our numerical integration we use only the values of $f(x)$ at certain points, namely, at x_1, \dots, x_q . It may even happen that $f(x)$ is not given by a formula at all, but is an empirical function given by those values which are obtained from an experiment or in some other way. In this case we cannot use the integral (4) but may consider the sum

$$(9) \quad E = \sum_{m=1}^q [f(x_m) - F(x_m)]^2.$$

⁶ It can be shown that for such a function f even the equality sign in (8) holds. Proof in Ref. [E3].

We may then ask for what choice of the α_m and β_m in (3) this quantity E becomes minimum. Necessary conditions are the $2N + 1$ conditions

$$(10) \quad \frac{\partial E}{\partial \alpha_0} = 0, \quad \frac{\partial E}{\partial \alpha_n} = 0, \quad \frac{\partial E}{\partial \beta_n} = 0, \quad (n = 1, \dots, N).$$

$F(x)$ should approximate the $q + 1$ given values $f(x_0), \dots, f(x_q)$ as closely as possible. But if $2N + 1 > q + 1$, then more than $q + 1$ constants appear in $F(x)$, there are an infinite number of ways that $F(x)$ can take on these $q + 1$ values, and the problem is meaningless. Therefore, we assume that $2N + 1 = q + 1$ or $N = q/2$ where q is even. Then it can be shown that (10) leads to $\alpha_0 = a_0$, $\alpha_n = a_n$, $\beta_n = b_n$ ($n = 1, 2, \dots, N - 1$), as given by (1), and

$$(11) \quad a_N = \frac{1}{q} \sum_{m=1}^q (-1)^m f(x_m) \quad \left(N = \frac{q}{2} \right).$$

For this choice of the coefficients in (3) the approximating trigonometric polynomial $F(x)$ assumes the values $F(x_m) = f(x_m)$, and the error E in (9) is zero. The proofs are somewhat long and can be found in Ref. [A11] or [A15] in Appendix 1.

Note that a_N obtained from (1) equals twice the value given by (11), as the student may verify, and $b_N = 0$ in (1) because when $n = N = q/2$ all the terms in the formula for b_n are zero.

For example, if $q = 24$, then $N = 12$, and $a_0, \dots, a_{11}, b_1, \dots, b_{11}$ should be computed from (1) while a_{12} should be computed from (11).

Problems

1. Carry out the details of the computations in Ex. 1.
2. Let f be an odd empirical function having period 2π and the following values in the interval $0 \leq x \leq \pi$:

x	0	$\frac{\pi}{12}$	$\frac{2\pi}{12}$	$\frac{3\pi}{12}$	$\frac{4\pi}{12}$	$\frac{5\pi}{12}$	$\frac{6\pi}{12}$	$\frac{7\pi}{12}$	$\frac{8\pi}{12}$	$\frac{9\pi}{12}$	$\frac{10\pi}{12}$	$\frac{11\pi}{12}$	
$f(x)$	0	0.20	0.36	0.53	0.70	0.84	0.93	0.98	1.00	0.97	0.85	0.50	0

Approximate $f(x)$ by a trigonometric polynomial $F(x)$ [cf. (3)] such that the error (9) is zero.

3. Let $f(x) = \begin{cases} -1 & \text{if } -\pi < x < 0 \\ 1 & \text{if } 0 < x < \pi \end{cases}$ and $f(x + 2\pi) = f(x)$. Find the function $F(x)$ of the form (3) for which the total square error (4) is minimum.
4. Compute the minimum square error in Prob. 3 for $N = 1, 3, 5, 7$. What is the smallest N such that $E^* \leq 0.2$?
5. Show that the minimum square error (7) is a monotone decreasing function of N .

In each case, find the function $F(x)$ of the form (3) for which the total square error E is minimum and compute this minimum value for $N = 1, 2, \dots, 5$.

6. $f(x) = x$ ($-\pi < x < \pi$)

7. $f(x) = x^2$ ($-\pi < x < \pi$)

Using (8) with the equality sign, show that:

8. $1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots = \frac{\pi^2}{8}$. Hint: consider $f(x) = \begin{cases} -1 & \text{if } -\pi < x < 0 \\ 1 & \text{if } 0 < x < \pi \end{cases}$

9. $1 + \frac{1}{2^4} + \frac{1}{3^4} + \cdots = \frac{\pi^4}{90}$. Hint: consider $f(x)$ in Prob. 6, Sec. 8.2.

10. Obtain the answer to Prob. 9 directly from the Fourier series of

$$f(x) = x^4 - 2\pi^2 x^2 \text{ when } -\pi < x < \pi \text{ and } f(x + 2\pi) = f(x).$$

8.9 Instrumental Methods for Determining Fourier Coefficients

In many cases it will be sufficient to determine numerical values of the first few Fourier coefficients of a given periodic function. For this purpose instrumental methods are of considerable help, in particular if the function is an empirical function, given by its graph or by numerical values, or if the corresponding integrals in the Euler formulas cannot be evaluated by elementary methods of integration. We shall discuss two suitable instruments, namely, the planimeter (Sec. 0.8) and the harmonic analyzer.

For obtaining numerical values of the first few Fourier coefficients of a given function $f(x)$ of period T by means of a planimeter, we have to plot graphs of the functions

$$\frac{1}{T}f(x), \quad \frac{2}{T}f(x)\cos\frac{2\pi x}{T}, \quad \frac{2}{T}f(x)\sin\frac{2\pi x}{T}, \quad \frac{2}{T}f(x)\cos\frac{4\pi x}{T}, \dots;$$

cf. (3) in Sec. 8.4, for $-T/2 \leq x \leq T/2$, and determine the areas under these curves.

This procedure can be simplified considerably by means of a harmonic analyzer which consists of a planimeter and an additional equipment. In Fig. 238 the planimeter (1) moves on a platform (2). The carriage (3) is moved on a rail (4) by means of a tracing arm (5), carrying the tracing pin (6). The gear (7) is exchangeable. The advantage of the harmonic analyzer is that for determining many Fourier coefficients we have to plot only the graph of the function $f(x)$. To obtain a certain Fourier coefficient of $f(x)$ we have to insert the appropriate gear and then trace around the complete boundary of the shaded area (Fig. 238), starting from the point A and returning to A ; the value of the Fourier coefficient is then obtained as the difference of the initial and final readings on the scale of the integrating wheel of the planimeter. To determine another Fourier coefficient of $f(x)$, we have to exchange the gear and proceed as before.⁷

⁷ The mechanical precision of a harmonic analyzer is of great importance for obtaining satisfactory results. The author has successfully used a Mader-Ott Harmonic Analyzer for many years.

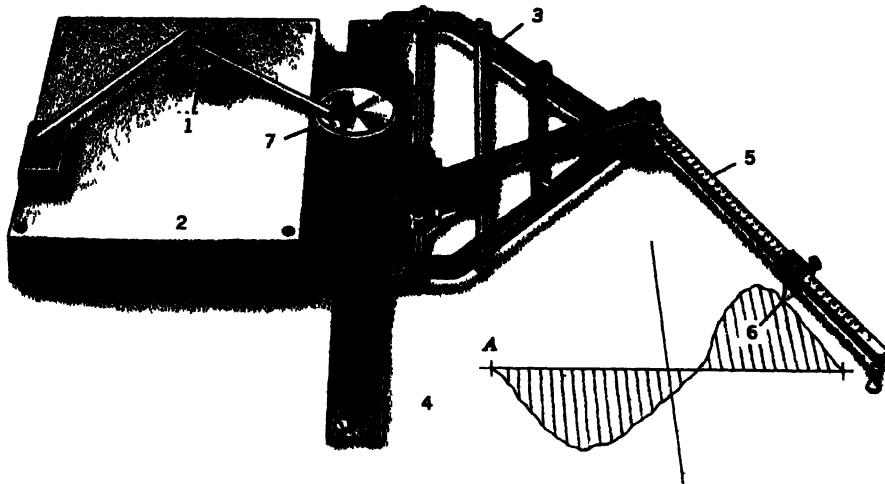


Fig. 238. Harmonic analyzer, manufactured by A. Ott, Kempten (Germany).

8.10 The Fourier Integral

Fourier series are powerful tools in treating various problems involving periodic functions. A first illustration of this fact was given in Sec. 8.7, and further important problems in connection with partial differential equations will be considered in Chap. 9. Since, of course, many practical problems do not involve periodic functions, it is desirable to generalize the method of Fourier series to include nonperiodic functions.

Roughly speaking, if we start with a periodic function $f_T(x)$ of period T and let T approach infinity, then the resulting function $f(x)$ is no longer periodic.

Example 1. Consider the function

$$f_T(x) = \begin{cases} 0 & \text{when } -T/2 < x < -1 \\ 1 & \text{when } -1 < x < 1 \\ 0 & \text{when } 1 < x < T/2 \end{cases}$$

having period $T > 2$ (cf. Fig. 239 on the next page). For $T \rightarrow \infty$ we obtain the function

$$f(x) = \lim_{T \rightarrow \infty} f_T(x) = \begin{cases} 1 & \text{when } -1 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Example 2. Let

$$f_T(x) = e^{-|x|} \quad \text{when } -T/2 < x < T/2 \quad \text{and} \quad f_T(x + T) = f_T(x).$$

Then (cf. Fig. 240 on the next page)

$$f(x) = \lim_{T \rightarrow \infty} f_T(x) = e^{-|x|}.$$

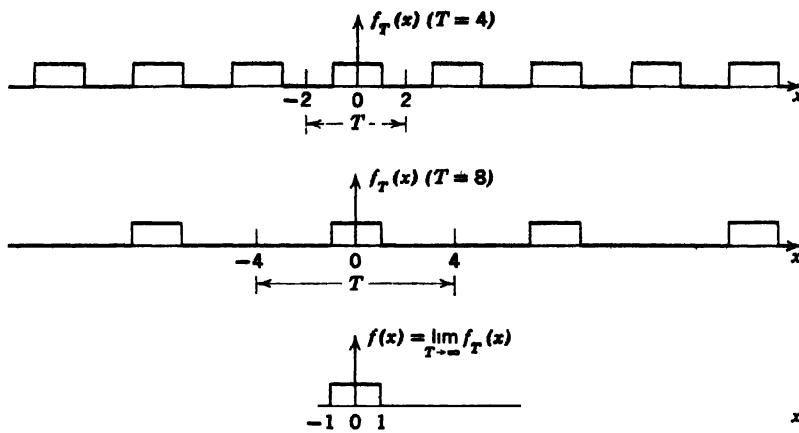


Fig. 239. Example 1.

Now let us start from a periodic function $f_T(x)$ that has period T and can be represented by a Fourier series:

$$f_T(x) = a_0 + \sum_{n=1}^{\infty} \left(a_n \cos \frac{2n\pi}{T} x + b_n \sin \frac{2n\pi}{T} x \right).$$

If we use the short notation

$$w_n = \frac{2n\pi}{T}$$

and insert a_n and b_n according to the Euler formulas (3) in Sec. 8.4, denoting the variable of integration by v , we obtain

$$f_T(x) = \frac{1}{T} \int_{-T/2}^{T/2} f_T(v) dv$$

$$\frac{2}{T} \sum_{n=1}^{\infty} \left[\cos w_n x \int_{-T/2}^{T/2} f_T(v) \cos w_n v dv + \sin w_n x \int_{-T/2}^{T/2} f_T(v) \sin w_n v dv \right].$$

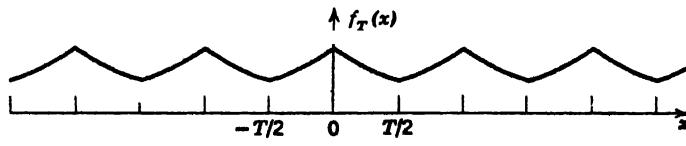


Fig. 240. Example 2.

Now,

$$w_{n+1} - w_n = \frac{2(n+1)\pi}{T} - \frac{2n\pi}{T} = \frac{2\pi}{T},$$

and we set

$$\Delta w = w_{n+1} - w_n = \frac{2\pi}{T}.$$

Then $2/T = \Delta w/\pi$, and we may write that Fourier series in the form

$$(1) \quad f_T(x) = \frac{1}{T} \int_{-T/2}^{T/2} f_T(v) dv + \frac{1}{\pi} \sum_{n=1}^{\infty} \left[\cos(w_n x) \Delta w \int_{-T/2}^{T/2} f_T(v) \cos w_n v dv + \sin(w_n x) \Delta w \int_{-T/2}^{T/2} f_T(v) \sin w_n v dv \right].$$

This representation is valid for any fixed T , arbitrarily large, but finite.

We now let T approach infinity and assume that the resulting nonperiodic function

$$f(x) = \lim_{T \rightarrow \infty} f_T(x)$$

is absolutely integrable on the x -axis, that is, the integral

$$(2) \quad \int_{-\infty}^{\infty} |f(x)| dx$$

exists. Then $1/T \rightarrow 0$ and the value of the first term on the right side of (1) approaches zero. Furthermore, $\Delta w = 2\pi/T \rightarrow 0$ and it seems plausible that the infinite series in (1) becomes an integral from 0 to ∞ , which represents $f(x)$, namely,

$$(3) \quad f(x) = \frac{1}{\pi} \int_0^{\infty} \left[\cos wx \int_{-\infty}^{\infty} f(v) \cos wv dv + \sin wx \int_{-\infty}^{\infty} f(v) \sin wv dv \right] dw.$$

If we introduce the short notations

$$(4) \quad A(w) = \int_{-\infty}^{\infty} f(v) \cos wv dv, \quad B(w) = \int_{-\infty}^{\infty} f(v) \sin wv dv,$$

this may be written in the form

$$(5) \quad f(x) = \frac{1}{\pi} \int_0^{\infty} [A(w) \cos wx + B(w) \sin wx] dw.$$

This is a representation of $f(x)$ by a so-called Fourier integral.

It is clear that our naive approach merely suggests the representation (5), but by no means establishes it; in fact, the limit of the series in (1) as Δw approaches zero is not the definition of the integral (3). Sufficient conditions for the validity of (5) are as follows.

Theorem 1. *If $f(x)$ is piecewise continuous (cf. Sec. 4.1) in every finite interval and has a right- and left-hand derivative at every point (cf. Sec. 8.2) and the*

integral (2) exists, then $f(x)$ can be represented by a Fourier integral. At a point where $f(x)$ is discontinuous the value of the Fourier integral equals the average of the left- and right-hand limits of $f(x)$ at that point (cf. Sec. 8.2). (Proof in Ref. [E2]; cf. Appendix 1.)

Example 3 (Single pulse, sine integral). Find the Fourier integral representation of the function (Fig. 241)

$$f(x) = \begin{cases} 1 & \text{when } |x| < 1, \\ 0 & \text{when } |x| > 1. \end{cases}$$

From (4) we obtain

$$A(w) = \int_{-\infty}^{\infty} f(v) \cos wv dv = \int_{-1}^1 \cos wv dv = \frac{\sin wv}{w} \Big|_{-1}^1 = \frac{2 \sin w}{w},$$

$$B(w) = \int_{-1}^1 \sin wv dv = 0,$$

and (5) becomes

$$(6) \quad f(x) = \frac{2}{\pi} \int_0^{\infty} \frac{\cos wx \sin w}{w} dw.$$

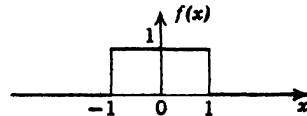


Fig. 241. Example 3.

The average of the left- and right-hand limits of $f(x)$ at $x = 1$ is equal to $(1 + 0)/2$, that is, $1/2$. Hence, from (6) and Theorem 1 we obtain the desired answer

$$\int_0^{\infty} \frac{\cos wx \sin w}{w} dw = \begin{cases} \pi/2 & \text{when } 0 \leq x < 1, \\ \pi/4 & \text{when } x = 1, \\ 0 & \text{when } x > 1. \end{cases}$$

We mention that this integral is called Dirichlet's⁸ discontinuous factor. Let us consider the case $x = 0$, which is of particular interest. When $x = 0$, then

$$(7) \quad \int_0^{\infty} \frac{\sin w}{w} dw = \frac{\pi}{2}.$$

We see that this integral is the limit of the so called sine integral

$$(8) \quad \text{Si}(z) = \int_0^z \frac{\sin w}{w} dw$$

as $z \rightarrow \infty$ (z real). The graph of $\text{Si}(z)$ is shown in Fig. 242.

In the case of a Fourier series the graphs of the partial sums are approximation curves of the curve of the periodic function represented by the series. Similarly, in the case of the Fourier integral (5), approximations are obtained by replacing ∞ by numbers a . Hence the integral

$$(9) \quad \int_0^a \frac{\cos wx \sin w}{w} dw$$

approximates the integral in (6) and therefore $f(x)$. Figure 243 shows oscillations near the points of discontinuity of $f(x)$.

⁸PETER GUSTAV LEJEUNE DIRICHLET (1805–1859), German mathematician, known by his important research work on Fourier series and in number theory.

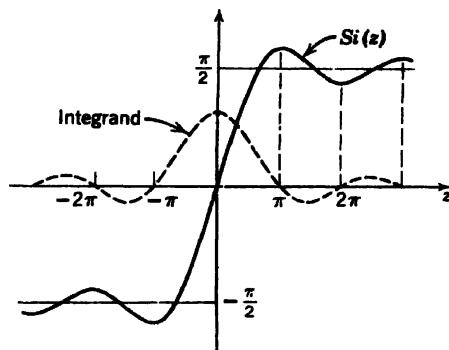
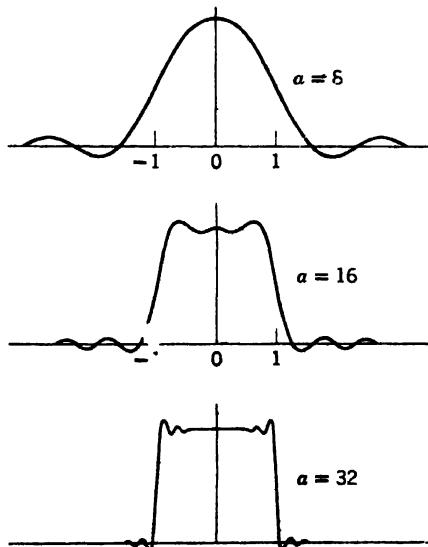


Fig. 242. Sine integral.

We might expect that these oscillations disappear as a approaches infinity, but this is not true; with increasing a , they are shifted closer to the points $x = \pm 1$. This unexpected behavior, which also occurs in connection with Fourier series, is known as the **Gibbs⁹ phenomenon**. It can be explained by representing (9) in terms of the sine integral as follows. Using (11b) in Sec. 0.1 we first have

$$\frac{2}{\pi} \int_0^a \frac{\cos wx \sin w}{w} dw = \frac{1}{\pi} \int_0^a \frac{\sin(w + wx)}{w} dw + \frac{1}{\pi} \int_0^a \frac{\sin(w - wx)}{w} dw.$$

In the first integral on the right we set $w + wx = t$. Then $dw/w = dt/t$, and $0 \leq w \leq a$ corresponds to $0 \leq t \leq (x+1)a$. In the last integral we set $w - wx = -t$. Then

Fig. 243. The integral (9) for $a = 8, 16$, and 32 .

⁹JOSIAH WILLARD GIBBS (1839–1903), American mathematician, whose work was of great importance to the development of vector analysis and mathematical physics.

$dw/w = dt/t$, and the interval $0 \leq w \leq a$ corresponds to $0 \leq t \leq (x-1)a$. Since $\sin(-t) = -\sin t$, we thus obtain

$$\frac{2}{\pi} \int_0^a \frac{\cos wx \sin w}{w} dw = \frac{1}{\pi} \int_0^{(x+1)a} \frac{\sin t}{t} dt - \frac{1}{\pi} \int_0^{(x-1)a} \frac{\sin t}{t} dt.$$

From this and (8) we see that our integral equals

$$\frac{1}{\pi} \text{Si}(a[x+1]) - \frac{1}{\pi} \text{Si}(a[x-1]),$$

and the oscillations in Fig. 243 result from those in Fig. 242. The increase of a amounts to a transformation of the scale on the axis and causes the shift of the oscillations. For further details see Ref. [E1] in Appendix 1.

If $f(x)$ is an even function, then $B(w) = 0$ in (4),

$$(10) \quad A(w) = 2 \int_0^\infty f(v) \cos wv dv,$$

and (5) reduces to the simpler form

$$(11) \quad f(x) = \frac{1}{\pi} \int_0^\infty A(w) \cos wx dw \quad (f \text{ even}).$$

If $f(x)$ is odd, then $A(w) = 0$,

$$(12) \quad B(w) = 2 \int_0^\infty f(v) \sin wv dv,$$

and (5) becomes

$$(13) \quad f(x) = \frac{1}{\pi} \int_0^\infty B(w) \sin wx dw \quad (f \text{ odd}).$$

These simplifications are quite similar to those in the case of a Fourier series discussed in Sec. 8.3.

Example 4 (Laplace integrals). Find the Fourier integral of

$$f(x) = e^{-kx} \quad \text{when } x > 0 \quad \text{and} \quad f(-x) = f(x) \quad (k > 0)$$

(cf. Fig. 240 in this section, where $k = 1$). Since f is even, we have from (10)

$$A(w) = 2 \int_0^\infty e^{-kv} \cos wv dv.$$

Now, by integration by parts,

$$\int e^{-kv} \cos wv dv = -\frac{k}{k^2 + w^2} e^{-kv} \left(-\frac{w}{k} \sin wv + \cos wv \right)$$

When $v = 0$, the expression on the right equals $-k/(k^2 + w^2)$; when v approaches infinity, it approaches zero because of the exponential factor. Thus

$$A(w) = \frac{2k}{k^2 + w^2},$$

and by substituting this in (11) we obtain the representation

$$f(x) = e^{-kx} = \frac{2k}{\pi} \int_0^\infty \frac{\cos wx}{k^2 + w^2} dw \quad (x > 0, k > 0).$$

From this representation we see that

$$(14) \quad \int_0^\infty \frac{\cos wx}{k^2 + w^2} dw = \frac{\pi}{2k} e^{-kx} \quad (x > 0, k > 0).$$

Similarly, from the Fourier integral (13) of the odd function

$$f(x) = e^{-kx} \quad \text{when } x > 0 \quad \text{and} \quad f(-x) = -f(x), \quad (k > 0)$$

we obtain the result

$$(15) \quad \int_0^\infty \frac{w \sin wx}{k^2 + w^2} dw = \frac{\pi}{2} e^{-kx} \quad (x > 0, k > 0).$$

The integrals (14) and (15) are the so-called *Laplace integrals*.

This example illustrates that the Fourier integral representation may be used for evaluating integrals.

Problems

Using the Fourier integral representation, show that

$$1. \int_0^\infty \frac{\cos xw + w \sin xw}{1 + w^2} dw = \begin{cases} 0 & \text{when } x < 0 \\ \pi/2 & \text{when } x = 0 \\ \pi e^{-x} & \text{when } x > 0 \end{cases} \quad [\text{Use (5).}]$$

$$2. \int_0^\infty \frac{1 - \cos \pi w}{w} \sin xw dw = \begin{cases} \frac{\pi}{2} & \text{when } 0 < x < \pi \\ 0 & \text{when } x > \pi \end{cases} \quad [\text{Use (13).}]$$

$$3. \int_0^\infty \frac{\cos xw}{1 + w^2} dw = \frac{\pi}{2} e^{-x} \quad (x > 0) \quad [\text{Use (11).}]$$

$$\int_0^\infty \frac{\sin w \cos xw}{w} dw = \begin{cases} \pi/2 & \text{when } 0 \leq x < 1 \\ \pi/4 & \text{when } x = 1 \\ 0 & \text{when } x > 1 \end{cases} \quad [\text{Use (11).}]$$

$$5. \int_0^\infty \frac{\cos(\pi w/2) \cos xw}{1 - w^2} dw = \begin{cases} \frac{\pi}{2} \cos x & \text{when } |x| < \frac{\pi}{2} \\ 0 & \text{when } |x| > \frac{\pi}{2} \end{cases} \quad [\text{Use (11).}]$$

$$6. \int_0^\infty \frac{w^3 \sin xw}{w^4 + 4} dw = \frac{\pi}{2} e^{-x} \cos x \quad \text{when } x > 0 \quad [\text{Use (13).}]$$

$$7. \int_0^\infty \frac{\sin \pi w \sin xw}{1 - w^2} dw = \begin{cases} \frac{\pi}{2} \sin x & \text{when } 0 \leq x \leq \pi \\ 0 & \text{when } x > \pi \end{cases} \quad [\text{Use (13).}]$$

Represent the following functions $f(x)$ in the form (11).

$$8. f(x) = \begin{cases} 1 & \text{when } 0 < x < a \\ 0 & \text{when } x > a \end{cases} \quad 9. f(x) = \begin{cases} x & \text{when } 0 < x < a \\ 0 & \text{when } x > a \end{cases}$$

$$10. f(x) = \frac{1}{1 + x^2} \quad [\text{cf. (14)}] \quad 11. f(x) = e^{-x} + e^{-2x} \quad (x > 0)$$

$$12. f(x) = \begin{cases} x & \text{when } 0 < x < 1 \\ 2-x & \text{when } 1 < x < 2 \\ 0 & \text{when } x > 2 \end{cases} \quad 13. f(x) = \begin{cases} x^2 & \text{when } 0 < x < a \\ 0 & \text{when } x > a \end{cases}$$

14. Show that $f(x) = 1$ ($0 < x < \infty$) cannot be represented by a Fourier integral.

If $f(x)$ has the representation (11), show that

$$15. f(ax) = \frac{1}{\pi a} \int_0^\infty A\left(\frac{w}{a}\right) \cos xw dw \quad (a > 0)$$

$$16. x^2 f(x) = \frac{1}{\pi} \int_0^\infty A^*(w) \cos xw dw, \quad A^* = -\frac{d^2 A}{dw^2}$$

17. Solve Prob. 13 by applying the formula in Prob. 16 to the result of Prob. 8.

18. Show that $x f(x) = \frac{1}{\pi} \int_0^\infty B^*(w) \sin xw dw$ where $B^* = -\frac{dA}{dw}$ and A is given by (10).

19. Verify the formula in Prob. 18 for $f(x) = 1$ when $0 < x < a$ and $f(x) = 0$ when $x > a$.

20. (Complex form of the Fourier integral, Fourier transform) Using (8), Sec. 0.1, show that (5) can be written

$$(16) \quad f(x) = \frac{1}{\pi} \int_0^\infty \left[\int_{-\infty}^\infty f(v) \cos(wx - wv) dv \right] dw.$$

Show that the integral from $-\infty$ to ∞ in (16) is an even function of w and (16) may be written

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^\infty \left[\int_{-\infty}^\infty f(v) \cos(wx - wv) dv \right] dw.$$

Show that

$$(17) \quad \frac{i}{2\pi} \int_{-\infty}^\infty \left[\int_{-\infty}^\infty f(v) \sin(wx - wv) dv \right] dw = 0,$$

so that by addition,

$$(18) \quad f(x) = \frac{1}{2\pi} \int_{-\infty}^\infty \left[\int_{-\infty}^\infty f(v) e^{iwx(x-v)} dv \right] dw.$$

This is the so-called *complex form* of the Fourier integral. Show that (18) can be written

$$(19) \quad f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty C(w) e^{iwx} dw$$

where

$$C(w) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty f(v) e^{-ivw} dv.$$

$f(x)$ is called the *Fourier transform* of (Cw) , and $C(w)$ is called the *inverse Fourier transform* of $f(x)$. (Elaborate tables of Fourier transforms are listed in Part E of Appendix 1.)

8.11 Orthogonal Functions

Let $g_m(x)$ and $g_n(x)$ be two real functions which are defined on an interval $a \leq x \leq b$ and are such that the integral of the product $g_m(x)g_n(x)$ over that interval exists. We shall denote this integral by (g_m, g_n) . This is a simple standard notation which is widely used. Thus

$$(1) \quad (g_m, g_n) = \int_a^b g_m(x)g_n(x) dx.$$

The functions are said to be **orthogonal on the interval $a \leq x \leq b$** if the integral (1) is zero, that is,

$$(2) \quad (g_m, g_n) = \int_a^b g_m(x)g_n(x) dx = 0 \quad (m \neq n).$$

A set of real functions $g_1(x), g_2(x), g_3(x), \dots$ is called an **orthogonal set of functions on an interval $a \leq x \leq b$** if these functions are defined on that interval and if all the integrals (g_m, g_n) exist and are zero for all pairs of distinct functions in the set.

The non-negative square root of (g_m, g_m) is called the **norm** of $g_m(x)$ and is generally denoted by $\|g_m\|$; thus

$$(3) \quad \|g_m\| = \sqrt{(g_m, g_m)} = \sqrt{\int_a^b g_m^2(x) dx}.$$

Throughout our consideration we shall make the following

General assumption. All occurring functions are bounded and are such that the occurring integrals exist, and the norms are not zero.

Clearly, an orthogonal set g_1, g_2, \dots on an interval $a \leq x \leq b$ whose functions have norm 1 satisfies the relations

$$(4) \quad (g_m, g_n) = \int_a^b g_m(x)g_n(x) dx = \begin{cases} 0 & \text{when } m \neq n \\ 1 & \text{when } m = n \end{cases} \quad m = 1, 2, \dots, \quad n = 1, 2, \dots.$$

Such a set is called an orthonormal set of functions on the interval $a \leq x \leq b$.

Obviously, from an orthogonal set, we may obtain an orthonormal set by dividing each function by its norm on the interval under consideration.

Example 1. The functions $g_m(x) = \sin mx$, $m = 1, 2, \dots$ form an orthogonal set on the interval $-\pi \leq x \leq \pi$, because

$$(5) \quad (g_m, g_n) = \int_{-\pi}^{\pi} \sin mx \sin nx dx = 0 \quad (m \neq n).$$

The norm $\|g_m\|$ equals $\sqrt{\pi}$, because

$$\|g_m\|^2 = \int_{-\pi}^{\pi} \sin^2 mx dx = \pi \quad (m = 1, 2, \dots).$$

Hence the corresponding orthonormal set consists of the functions

$$\frac{\sin x}{\sqrt{\pi}}, \quad \frac{\sin 2x}{\sqrt{\pi}}, \quad \frac{\sin 3x}{\sqrt{\pi}}, \quad \dots$$

Example 2. The functions

$$1, \quad \cos x, \quad \sin x, \quad \cos 2x, \quad \sin 2x, \quad \dots$$

occurring in Fourier series form an orthogonal set on the interval $-\pi \leq x \leq \pi$. This follows from (5), the analogous relation for the cosine functions, and

$$\int_{-\pi}^{\pi} \cos mx \sin nx dx = 0 \quad \text{for all } m, n = 0, 1, \dots$$

The corresponding orthonormal set is

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{\cos x}{\sqrt{\pi}}, \quad \frac{\sin x}{\sqrt{\pi}}, \quad \frac{\cos 2x}{\sqrt{\pi}}, \quad \frac{\sin 2x}{\sqrt{\pi}}, \quad \dots$$

Looking at the derivation of the Euler formulas (6) in Sec. 8.2 for the Fourier coefficients, we see that we used merely the fact that the set in Ex. 2, this section, is orthogonal on an interval of length 2π . This simple observation suggests the attempt to represent given functions $f(x)$ in terms of any other orthogonal set $g_1(x), g_2(x), \dots$ in the form

$$(6) \quad f(x) = \sum_{n=1}^{\infty} c_n g_n(x) = c_1 g_1(x) + c_2 g_2(x) + \dots$$

and determine the coefficients c_1, c_2, \dots as in Sec. 8.2. If the series (6) converges and represents $f(x)$, it is called a **generalized Fourier series** of $f(x)$, and its coefficients are called the **Fourier constants of $f(x)$ with respect to that orthogonal set of functions**.

To determine these constants, we proceed as in Sec. 8.2. We multiply both sides of (6) by $g_m(x)$ and integrate over the interval $a \leq x \leq b$ on which the functions are orthogonal; assuming that term-by-term integration is permissible (cf. footnote 2 in Sec. 8.2), we obtain

$$\int_a^b f g_m dx = \sum_{n=1}^{\infty} c_n \int_a^b g_n g_m dx.$$

The integral for which $n = m$ is equal to the square of the norm $\|g_m\|$, while all the other integrals are zero because the functions are orthogonal. Thus,

$$(7') \quad \int_a^b f g_m dx = c_m \|g_m\|^2$$

and the desired formula for the Fourier constants is

$$(7) \quad c_m = \frac{1}{\|g_m\|^2} \int_a^b f(x) g_m(x) dx.$$

Note that for the functions in Ex. 2 this formula becomes identical with the Euler formulas (6) in Sec. 8.2.

As in Sec. 8.8 we conclude that if the set of functions is orthonormal, those Fourier constants satisfy the Bessel inequality

$$(8) \quad c_1^2 + c_2^2 + c_3^2 + \cdots \leq \int_a^b f^2(x) dx.$$

Hence the series on the left converges and, therefore,

$$c_n \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty.$$

Some important sets of real functions g_1, g_2, \dots occurring in applications are not orthogonal but have the property that for some function $p(x)$,

$$(9) \quad \int_a^b p(x)g_m(x)g_n(x) dx = 0 \quad \text{when } m \neq n.$$

Such a set is then said to be *orthogonal with respect to the weight function $p(x)$ on the interval $a \leq x \leq b$* . The *norm* of g_m is now defined as

$$(10) \quad \|g_m\| = \sqrt{\int_a^b p(x)g_m^2(x) dx},$$

and if the norm of each function g_m is 1, the set is said to be *orthonormal* on that interval with respect to $p(x)$.

If we set $h_m = \sqrt{p} g_m$, then (9) becomes

$$\int_a^b h_m(x)h_n(x) dx = 0 \quad (m \neq n),$$

that is, the functions h_m form an orthogonal set in the usual sense. Clearly, if these functions are to be real, the weight function must be non-negative. Important examples will be considered in the following sections.

Problems

1. Show that the integral in (1) has the properties

$$(g_m, g_n) = (g_n, g_m), \quad (ag_k + bg_m, g_n) = a(g_k, g_n) + b(g_m, g_n).$$

2. Prove (8).

In each case, show that the given set is orthogonal on the given interval I and determine the corresponding orthonormal set.

- ✓ 3. 1, $\cos x, \cos 2x, \cos 3x, \dots$, $I: -\pi \leq x \leq \pi$
- 4. 1, $\cos x, \cos 2x, \cos 3x, \dots$, $I: 0 \leq x \leq 2\pi$
- ✓ 5. 1, $\cos 2x, \cos 4x, \cos 6x, \dots$, $I: 0 \leq x \leq \pi$
- 6. $\sin mx, \sin 2mx, \sin 3mx, \dots$, $I: -1 \leq x \leq 1$
- ✓ 7. 1, $\cos mx, \sin mx, \cos 2mx, \sin 2mx, \dots$, $I: 0 \leq x \leq 2$
- 8. 1, $\cos \frac{2n\pi}{T} x$ ($n = 1, 2, \dots$), $I: 0 \leq x \leq T$
- 9. 1, $\cos 2x, \sin 2x, \cos 4x, \sin 4x, \dots$, $I: -\pi/2 \leq x \leq \pi/2$

10. $1, \cos \frac{2n\pi}{T}x, \sin \frac{2n\pi}{T}x$ ($n = 1, 2, \dots$). $I: -\frac{T}{2} \leq x \leq \frac{T}{2}$
- ✓ 11. $1, x, (3x^2 - 1)/2, (5x^2 - 3x)/2$, $I: -1 \leq x \leq 1$
12. Show that for the functions in Prob. 10 the formula (7) becomes identical with the Euler formulas (3) in Sec. 8.4.
13. Show that if the functions $g_1(x), g_2(x), \dots$ form an orthogonal set on the interval $a \leq x \leq b$, then the functions $g_1(cx + k), g_2(cx + k), \dots$, $c \neq 0$, form an orthogonal set on the interval $(a - k)/c \leq t \leq (b - k)/c$.
14. Using the result of Prob. 13, derive the orthogonality property of the set in Prob. 6 from that in Ex. 1.
- ✓ 15. Determine constants a_0, b_0, \dots, c_3 so that the functions $g_1 = a_0, g_2 = b_0 + b_1x, g_3 = c_0 + c_1x + c_2x^2$ form an orthonormal set on the interval $-1 \leq x \leq 1$.
16. Determine the constants in Prob. 15 so that the functions $g_1(x), g_2(x), g_3(x)$ form an orthonormal set on the interval $0 \leq x \leq 1$.
17. (Legendre polynomials) The Legendre polynomials $P_n(x)$, $n = 0, 1, \dots$ [cf. (11') in Sec. 3.3] form an orthogonal set on the interval $-1 \leq x \leq 1$. Using Prob. 11 in Sec. 3.3, find the corresponding orthonormal set. Show that $g_1 = P_0/\|P_0\|$, $g_2 = P_1/\|P_1\|$, and $g_3 = P_2/\|P_2\|$, where g_1, g_2, g_3 are the functions obtained in Prob. 15.
18. Let $S_N(x)$ be the N th partial sum of (6). Then if

$$\lim_{N \rightarrow \infty} \int_a^b [f(x) - S_N(x)]^2 dx = 0,$$

the sum $S_N(x)$ is said to converge in the mean to the function $f(x)$. If this relation holds for each function $f(x)$ that satisfies the general assumption (p. 473), the set $g_1(x), g_2(x), \dots$ is said to be closed. Show that for a closed orthonormal set,

$$\sum_{n=1}^{\infty} c_n^2 = \int_a^b f^2(x) dx \quad (\text{Parseval's formula}).$$

Hint: insert $S_N = \sum_{n=1}^N c_n g_n(x)$ into the above integral, use (7'), and show that

$$\int_a^b \left(\sum_{n=1}^N c_n g_n \right)^2 dx = \sum_{n=1}^N c_n^2. \text{ Then let } N \rightarrow \infty.$$

19. Show that a closed orthonormal set of functions is complete, that is, there is no function $f(x)$ that satisfies the general assumption on p. 473 and is orthogonal to each function of the set on the interval under consideration.
20. Find an (unbounded) function $f(x)$ such that $\int_0^1 f(x) dx$ exists, but $\int_0^1 f^2(x) dx$ does not exist.

8.12 Sturm-Liouville Problem

Various important orthogonal sets of functions arise as solutions of second-order differential equations of the form

$$(1) \quad [r(x)y']' + [q(x) + \lambda p(x)]y = 0$$

on some interval $a \leq x \leq b$, satisfying conditions of the form

$$(2) \quad (\text{a}) \quad k_1y + k_2y' = 0 \quad \text{at } x = a \quad (\text{b}) \quad l_1y + l_2y' = 0 \quad \text{at } x = b;$$

here λ is a parameter, and k_1, k_2, l_1, l_2 are given real constants, at least one in each condition (2) being different from zero.

The equation (1) is known as the **Sturm-Liouville¹⁰ equation**, and we shall see that Legendre's equation, Bessel's equation, and other important equations can be written in the form (1). The conditions (2) refer to the boundary points $x = a$ and $x = b$ of that interval and, therefore, are called **boundary conditions**. A differential equation together with boundary conditions constitutes what is known as a **boundary value problem**. Our boundary value problem (1), (2) is called a *Sturm-Liouville problem*.

Of course, this problem has the trivial solution $y \equiv 0$ for any value of the parameter λ . Solutions $y \neq 0$ are called *characteristic functions* or *eigenfunctions* of the problem, and the values of λ for which such solutions exist, are called *characteristic values* or *eigenvalues* of the problem.

Example 1. Find the eigenvalues and eigenfunctions of the Sturm-Liouville problem

$$(3) \quad (\text{a}) \quad y' + \lambda y = 0 \quad (\text{b}) \quad y(0) = 0, \quad y(\pi) = 0.$$

For negative $\lambda = -r^2$ the general solution of the equation is

$$y(x) = c_1 e^{rx} + c_2 e^{-rx}. \quad \checkmark$$

From (3b), $c_1 = c_2 = 0$ and $y \equiv 0$, which is not an eigenfunction. For $\lambda = 0$ the situation is similar. For positive $\lambda = r^2$ the general solution is

$$y(x) = A \cos rx + B \sin rx.$$

From the boundary conditions we obtain $y(0) = A = 0$ and thus,

$$y(\pi) = B \sin r\pi = 0 \quad \text{or} \quad r = 0, \pm 1, \pm 2, \dots$$

For $r = 0$ we have $y \equiv 0$. For $\lambda = r^2 = 1, 4, 9, 16, \dots$, taking $B = 1$, we obtain

$$y(x) = \sin rx \quad r = 1, 2, \dots$$

These are the eigenfunctions of the problem, and the eigenvalues are $\lambda = r^2$ where $r = 1, 2, \dots$.

It can be shown that under rather general conditions on the functions p , q , and r in (1) the Sturm-Liouville problem (1), (2) has infinitely many eigenvalues; the corresponding rather complicated theory can be found in Ref. [B5]; cf. Appendix 1.

Furthermore, the eigenfunctions are orthogonal, as follows.

Theorem 1. Let the functions p , q , and r in the Sturm-Liouville equation (1) be real and continuous on the interval $a \leq x \leq b$. Let $y_m(x)$ and $y_n(x)$ be eigenfunctions of the Sturm-Liouville problem (1), (2) corresponding to distinct eigenvalues λ_m and λ_n , respectively, and let the derivatives $y_m'(x)$ and $y_n'(x)$ be

¹⁰ JACQUES CHARLES FRANCOIS STURM (1803–1855), Swiss mathematician, JOSEPH LIOUVILLE (1809–1882), French mathematician, known by his important research work in complex analysis.

continuous on that interval. Then y_m and y_n are orthogonal on that interval with respect to the weight function p .

If $r(a) = 0$, then (2a) can be dropped from the problem. If $r(b) = 0$, then (2b) can be dropped. If $r(a) = r(b)$, then (2) can be replaced by

$$(4) \quad y(a) = y(b), \quad y'(a) = y'(b).$$

Proof. y_m satisfies

$$(ry_m')' + (q + \lambda_m p)y_m = 0,$$

and y_n satisfies

$$(ry_n')' + (q + \lambda_n p)y_n = 0.$$

Multiplying the first equation by y_n and the second by $-y_m$ and adding, we get

$$\begin{aligned} (\lambda_m - \lambda_n)py_m y_n &= y_m(ry_n')' - y_n(ry_m')' \\ &= [(ry_n')y_m - (ry_m')y_n]' \end{aligned}$$

where the last equality can be readily verified by carrying out the indicated differentiation in the last expression. Integrating over x from a to b , we obtain

$$(5) \quad (\lambda_m - \lambda_n) \int_a^b py_m y_n dx = \left[r(y_n'y_m - y_m'y_n) \right]_a^b.$$

The expression on the right equals

$$\begin{aligned} (6) \quad r(b)[y_n'(b)y_m(b) - y_m'(b)y_n(b)] \\ - r(a)[y_n'(a)y_m(a) - y_m'(a)y_n(a)]. \end{aligned}$$

Case 1. If $r(a) = 0$ and $r(b) = 0$, then the expression in (6) is zero. Hence the expression on the left-hand side of (5) must be zero, and since λ_m and λ_n are distinct, we obtain the desired orthogonality

$$(7) \quad \int_a^b p(x)y_m(x)y_n(x) dx = 0 \quad (m \neq n)$$

without the use of the boundary conditions (2).

Case 2. Let $r(b) = 0$, but $r(a) \neq 0$. Then the first line in (6) is zero. We consider the remaining expression in (6). From (2a) we have

$$\begin{aligned} k_1 y_n(a) + k_2 y_n'(a) &= 0, \\ k_1 y_m(a) + k_2 y_m'(a) &= 0. \end{aligned}$$

Let $k_2 \neq 0$. Then by multiplying the first equation by $y_m(a)$ and the last by $-y_n(a)$ and adding, we have

$$k_2[y_n'(a)y_m(a) - y_m'(a)y_n(a)] = 0.$$

Since $k_2 \neq 0$, the expression in brackets must be zero. This expression is identical with that in the last line of (6). Hence (6) is zero, and from (5) we see in (7). If $k_2 = 0$, then by assumption $k_1 \neq 0$, and the argument of proof is similar.

Case 3. If $r(a) = 0$, but $r(b) \neq 0$, the proof is similar to that in Case 2, but instead of (2a) we now have to use (2b).

Case 4. If $r(a) \neq 0$ and $r(b) \neq 0$, we have to use both boundary conditions (2) and proceed as in Cases 2 and 3.

Case 5. Let $r(a) = r(b)$. Then (6) takes the form

$$r(b)[y_n'(b)y_m(b) - y_m'(b)y_n(b) - y_n'(a)y_m(a) + y_m'(a)y_n(a)].$$

We may use (2) as before and conclude that the expression in brackets is zero. However, we immediately see that this would also follow from (4), so that we may replace (2) by (4). Hence (5) yields (7), as before. This completes the proof of Theorem 1.

Example 2. The differential equation in Ex. 1 is of the form (1) where $r = 1$, $q = 0$, and $p = 1$. From Theorem 1 it follows that the eigenfunctions are orthogonal on the interval $0 \leq x \leq \pi$.

Example 3. The reader may show that the functions

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \dots$$

occurring in the Fourier series of a function of period 2π are the eigenfunctions of the Sturm-Liouville problem

$$y'' + \lambda y = 0, \quad y(\pi) = y(-\pi), \quad y'(\pi) = y'(-\pi).$$

Hence from Theorem 1 it follows that these functions form an orthogonal set on the interval $-\pi \leq x \leq \pi$. Note that the boundary conditions of our present problem are of the form (4).

We shall now prove a basic theorem concerning the eigenvalues of Sturm-Liouville problems.

Theorem 2. If the Sturm-Liouville problem (1), (2) satisfies the condition stated in Theorem 1 and p is positive in the whole interval $a \leq x \leq b$ (or negative everywhere in that interval), then all the eigenvalues of the problem are real.

Proof. Let $\lambda = \alpha + i\beta$ be an eigenvalue of the problem and let

$$y(x) = u(x) + iv(x)$$

be a corresponding eigenfunction; here α , β , u , and v are real. Substituting this into (1) we have

$$(ru' + irv')' + (q + xp + i\beta p)(u + iv) = 0.$$

This complex equation is equivalent to the following pair of equations for the real and imaginary parts:

$$\begin{aligned} (ru')' + (q + xp)u - \beta pr &= 0, \\ (rv')' + (q + xp)v + \beta pu &= 0. \end{aligned}$$

Multiplying the first equation by v and the second by $-u$ and adding, we obtain

$$\begin{aligned} -\beta(u^2 + v^2)p &= u(rv')' - v(ru')' \\ &= [(rv')u - (ru')v]' . \end{aligned}$$

Integrating over x from a to b , we find

$$-\beta \int_a^b (u^2 + v^2)p \, dx = \left[r(uv' - u'v) \right]_a^b.$$

From the boundary conditions it follows that the expression on the right is zero; this can be shown in a fashion similar to that in the proof of Theorem 1. Since y is an eigenfunction, $u^2 + v^2 \not\equiv 0$. Since y and p are continuous and $p > 0$ (or $p < 0$) for all x between a and b , the integral on the left is not zero. Hence, $\beta = 0$, which means that $\lambda = \alpha$ is real. This completes the proof.

Theorem 2 is illustrated by Exs. 2, 3, and the following example.

Example 4 (Legendre polynomials). Legendre's equation (cf. Sec. 3.3) can be written

$$[(1 - x^2)y']' + \lambda y = 0, \quad \lambda = n(n + 1),$$

and is, therefore, a Sturm-Liouville equation (1) with $r = 1 - x^2$, $q = 0$, and $p = 1$. Since $r = 0$ when $x = \pm 1$, no boundary conditions are needed to form a Sturm-Liouville problem on the interval $-1 \leq x \leq 1$. We know that, for $n = 0, 1, \dots$, the Legendre polynomials $P_n(x)$ are solutions of the problem. Hence these are the eigenfunctions, and since they have continuous derivatives, it follows from Theorem 1 that they are orthogonal on that interval, that is,

$$(8) \quad \int_{-1}^1 P_m(x)P_n(x) \, dx = 0 \quad (m \neq n).$$

The norm is (cf. Prob. 11 in Sec. 3.3)

$$(9) \quad \|P_m\| = \sqrt{\int_1^1 P_m^2(x) \, dx} = \sqrt{\frac{2}{2m + 1}} \quad m = 0, 1, \dots$$

If $g_1(x), g_2(x), \dots$ are eigenfunctions which are orthogonal on the interval $a \leq x \leq b$ with respect to a weight function $p(x)$, and if a given function $f(x)$ can be represented by a generalized Fourier series

$$(10) \quad f(x) = c_1 g_1(x) + c_2 g_2(x) + \dots$$

involving these functions, then the Fourier constants c_1, c_2, \dots of $f(x)$ in this eigenfunction expansion can be determined as in Sec. 8.11; the only difference is that we have to multiply that series by pg_m instead of g_m before integrating. The remaining steps are the same as before and yield

$$(11) \quad c_m = \frac{1}{\|g_m\|^2} \int_a^b p(x)f(x)g_m(x) \, dx, \quad m = 1, 2, \dots,$$

where the norm $\|g_m\|$ is now defined by (10), Sec. 8.11.

Problems

Find the eigenfunctions of the following Sturm-Liouville problems.

1. $y'' + \lambda y = 0, \quad y(0) = 0, \quad y'(l) = 0$
2. $y'' + \lambda y = 0, \quad y'(0) = 0, \quad y'(l) = 0$

Find a Sturm-Liouville problem whose eigenfunctions are

3. $1, \cos, \cos 2x, \cos 3x, \dots$

4. $1, \cos \frac{2n\pi}{T}x, \sin \frac{2n\pi}{T}x \quad (n = 1, 2, \dots)$

5. (**Legendre polynomials**) Show that the functions $P_n(\cos \theta)$, $n = 0, 1, \dots$, form an orthogonal set on the interval $0 \leq \theta \leq \pi$ with respect to the weight function $p(\theta) = \sin \theta$.
6. Show that if $f(x)$ is even, then c_1, c_3, c_5, \dots in $f(x) = c_0 P_0 + c_1 P_1(x) + \dots$ are zero.
7. State the reason why the integral of $P_n(x)$ ($n = 1, 2, \dots$) from -1 to 1 is zero.

In each case obtain the first few terms of the representation of $f(x)$ by a generalized Fourier series involving Legendre polynomials and graph the first three partial sums.

8. $f(x) = \begin{cases} 0 & \text{if } -1 < x < 0 \\ 1 & \text{if } 0 < x < 1 \end{cases}$

9. $f(x) = \begin{cases} 0 & \text{if } -1 < x < 0 \\ x & \text{if } 0 < x < 1 \end{cases}$

10. $f(x) = |x| \quad \text{if } -1 < x < 1$

11. $f(x) = x^3$

12. (**Tchebichef¹¹ polynomials**) The functions

$$T_n(x) = \cos(n \arccos x), \quad U_n(x) = \sin(n \arccos x), \quad n = 0, 1, \dots$$

are called *Tchebichef polynomials of the first and second kind*, respectively. Show that

$$\begin{aligned} T_0 &= 1, & T_1(x) &= x, & T_2(x) &= 2x^2 - 1, \\ U_0 &= 0, & U_1(x) &= \sqrt{1-x^2}, & U_2(x) &= 2x\sqrt{1-x^2}. \end{aligned}$$

13. Show that the Tchebichef polynomials $T_n(x)$ are orthogonal on the interval $-1 \leq x \leq 1$ with respect to the weight function $p(x) = 1/\sqrt{1-x^2}$. Hint: to evaluate the integral set $\arccos x = \theta$.

14. Find $T_3(x)$ and $U_3(x)$. Show that $T_n(1) = 1$ and $U_n(1) = 0$.

15. (**Hermite¹² polynomials**) The functions

$$H_0 = 1, \quad H_n(x) = (-1)^n e^{x^2/2} \frac{d^n e^{-x^2/2}}{dx^n}, \quad n = 1, 2, \dots$$

are called *Hermite polynomials*. Show that

$$H_1(x) = x, \quad H_2(x) = x^2 - 1, \quad H_3(x) = x^3 - 3x, \quad H_4(x) = x^4 - 6x^2 + 3.$$

(As is true for many special functions, the literature contains more than one notation, and one often defines as Hermite polynomials the expressions

$$H_0^* = 1, \quad H_n^*(x) = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n}.$$

This differs from our definition which is preferably used in applications.)

16. Show that the Hermite polynomials satisfy the relation

$$H_{n+1}(x) = xH_n(x) - H_n'(x).$$

¹¹ PAFNUTI TCHEBICHEF (1821–1894), Russian mathematician, particularly known for his important contributions to the theory of numbers.

¹² CHARLES HERMITE (1822–1901), French mathematician, is known by his work in algebra and number theory.

17. Show that the Hermite polynomials are related to the coefficients of the Maclaurin series

$$e^{tx-t^2/2} = \sum_{n=0}^{\infty} a_n(x) t^n$$

by the formula $H_n(x) = n! a_n(x)$. Hint: note that $tx - t^2/2 = x^2/2 - (x - t)^2/2$. (The exponential function on the left is called the *generating function* of the H_n .)

18. Differentiating the generating function in Prob. 17 with respect to x , show that

$$H_n'(x) = n H_{n-1}(x).$$

19. Using the formula in Prob. 16 (with n replaced by $n - 1$) and the relation in Prob. 18 (and its derivative), prove that $H_n(x)$ satisfies the differential equation

$$y'' - xy' + ny = 0.$$

20. Solve the differential equation in Prob. 19 by the power series method.

21. Using the equation in Prob. 19, show that $w = e^{-x^2/4} H_n(x)$ is a solution of Weber's¹³ equation

$$w'' + (n + \frac{1}{2} - \frac{1}{4}x^2)w = 0 \quad (n = 0, 1, \dots).$$

22. Show that the Hermite polynomials are orthogonal on the x -axis $-\infty < x < \infty$ with respect to the weight function $p(x) = e^{-x^2/2}$. Hint. use the formulas in Probs. 15 and 18 and integrate by parts.

23. (Laguerre¹⁴ polynomials) The functions

$$L_0 = 1, \quad L_n(x) = \frac{e^x}{n!} \frac{d^n(x^n e^{-x})}{dx^n}, \quad n = 1, 2, \dots$$

are called *Laguerre polynomials*. Show that

$$L_1(x) = 1 - x, \quad L_2(x) = 1 - 2x + x^2/2, \quad L_3(x) = 1 - 3x + 3x^2/2 - x^3/6.$$

24. Show that

$$L_n(x) = \sum_{m=0}^n \frac{(-1)^m}{m!} \binom{n}{m} x^m = 1 - nx + \frac{n(n-1)}{4} x^2 - + \dots + \frac{(-1)^n}{n!} x^n.$$

25. $L_n(x)$ satisfies Laguerre's differential equation

$$xy'' + (1 - x)y' + ny = 0.$$

Verify this fact for $n = 0, 1, 2, 3$.

26. Verify by direct integration that $L_0, L_1(x), L_2(x)$ are orthogonal on the positive axis $0 \leq x < \infty$ with respect to the weight function $p(x) = e^{-x}$.

27. Prove that all the Laguerre polynomials have the orthogonality property in Prob.

26. Hint: consider $\int_0^\infty e^{-x} L_m(x) L_n(x) dx$ for $m < n$. Since the highest power in

- L_m is x^m , conclude that it suffices to show that $\int_0^\infty e^{-x} x^k L_n(x) dx = 0$ ($k < n$).

Prove this relation by repeated integration by parts.

28. Carry out the details of the proof of Theorem 1 in Cases 3 and 4.

Carry out the steps of the proof of Theorem 1 for the Sturm-Liouville problem in

29. Ex. 4

30. Ex. 1

¹³ HEINRICH WEBER (1842–1913), German mathematician.

¹⁴ EDMOND LAGUERRE (1834–1886), French mathematician who did research work in geometry and the theory of infinite series.

8.13 Orthogonality of Bessel Functions

We shall now consider the orthogonality of the Bessel functions J_n , which is quite important in applications, for example, in connection with vibrations of a circular membrane. We know that the functions $J_n(s)$ satisfy the Bessel equation (Sec. 3.5), that is,

$$s^2 \frac{d^2 J_n}{ds^2} + s \frac{dJ_n}{ds} + (s^2 - n^2) J_n = 0.$$

We assume that n is a nonnegative integer. Setting $s = \lambda x$ where λ is a constant, not zero, we have $dx/ds = 1/\lambda$, and

$$\frac{dJ_n}{ds} = \frac{dJ_n}{dx} \frac{dx}{ds} = \frac{J_n'}{\lambda}, \quad \frac{d^2 J_n}{ds^2} = \frac{J_n''}{\lambda^2},$$

where primes denote derivatives with respect to x . By substitution,

$$x^2 J_n''(\lambda x) + x J_n'(\lambda x) + (\lambda^2 x^2 - n^2) J_n(\lambda x) = 0.$$

Dividing by x , we may write this equation in the form

$$(1) \quad [x J_n'(\lambda x)]' + \left(-\frac{n^2}{x} + \lambda^2 x \right) J_n(\lambda x) = 0.$$

We see that for each fixed n this is a Sturm-Liouville equation (1), Sec. 8.12, with the parameter written as λ^2 instead of λ , and

$$p(x) = x, \quad q(x) = -n^2/x, \quad r(x) = x.$$

Since $r(x) = 0$ at $x = 0$, it follows from Theorem 1 in the last section that those solutions of (1) on a given interval $0 \leq x \leq R$ which satisfy the boundary condition

$$(2) \quad J_n(\lambda R) = 0$$

form an orthogonal set on that interval with respect to the weight function $p(x) = x$. (Note that for $n \neq 0$ the function q is discontinuous at $x = 0$, but this does not affect the proof of that theorem.) It can be shown (Ref. [B9]) that $J_n(s)$ has infinitely many real zeros; let $\alpha_{1n} < \alpha_{2n} < \alpha_{3n} \dots$ denote the positive zeros of $J_n(s)$. Then (2) holds for

$$(3) \quad \lambda R = \alpha_{mn} \quad \text{or} \quad \lambda = \lambda_{mn} = \frac{\alpha_{mn}}{R} \quad (m = 1, 2, \dots),$$

and since the derivative J_n' is continuous, also at $x = 0$, we obtain the following result.

Theorem 1. For each fixed $n = 0, 1, \dots$ the Bessel functions $J_n(\lambda_{1n}x)$, $J_n(\lambda_{2n}x)$, $J_n(\lambda_{3n}x)$, \dots , with λ_{mn} according to (3), form an orthogonal set on the interval $0 \leq x \leq R$ with respect to the weight function $p(x) = x$, that is,

$$(4) \quad \int_0^R x J_n(\lambda_{mn}x) J_n(\lambda_{kn}x) dx = 0 \quad (k \neq m).$$

Hence we have obtained infinitely many orthogonal sets, each corresponding to one of the fixed values of n .

For determining the Fourier constants of a given function with respect to such a set $J_n(\lambda_{mn}x)$ (n fixed, $m = 1, 2, \dots$), we must know the norms of these functions. The norms can be found as follows.

Multiplying (1) by $2xJ_n'(\lambda x)$, we may write the resulting equation in the form

$$\{[xJ_n'(\lambda x)]^2\}' + (\lambda^2 x^2 - n^2)\{J_n^2(\lambda x)\}' = 0.$$

Integrating over x from 0 to R we find

$$(5) \quad [xJ_n'(\lambda x)]^2 \Big|_0^R = - \int_0^R (\lambda^2 x^2 - n^2)\{J_n^2(\lambda x)\}' dx.$$

From (2') in Sec. 3.6, writing s and n instead of x and v , we have

$$s^{-n} \frac{dJ_n}{ds} = ns^{-n-1} J_n(s) - s^{-n} J_{n+1}(s).$$

Multiplying by s^{n+1} and setting $s = \lambda x$, we obtain

$$\lambda x \frac{dJ_n(\lambda x)}{d(\lambda x)} = x \frac{dJ_n(\lambda x)}{dx} = n J_n(\lambda x) - \lambda x J_{n+1}(\lambda x).$$

Hence the left-hand side of (5) equals

$$\left[[n J_n(\lambda x) - \lambda x J_{n+1}(\lambda x)]^2 \right]_{x=0}^R.$$

If $\lambda = \lambda_{mn}$, then $J_n(\lambda R) = 0$, and since $J_n(0) = 0$ ($n = 1, 2, \dots$), that left-hand side becomes

$$(6) \quad \lambda_{mn}^2 R^2 J_{n+1}^2(\lambda_{mn} R).$$

If we integrate by parts, the right-hand side of (5) becomes

$$-\left[(\lambda^2 x^2 + n^2) J_n^2(\lambda x) \right]_0^R + 2\lambda^2 \int_0^R x J_n^2(\lambda x) dx.$$

When $\lambda = \lambda_{mn}$, the first expression is zero, and from this and (6), dividing by λ_{mn}^2 , we obtain

$$(7) \quad \|J_n(\lambda_{mn}x)\|^2 = \int_0^R x J_n^2(\lambda_{mn}x) dx = \frac{R^2}{2} J_{n+1}^2(\lambda_{mn}R).$$

This is the square of the norm of the functions under consideration; here λ_{mn} is given by (3), and $\lambda_{mn}R = \alpha_{mn}$.

Consequently, if a function $f(x)$ can be represented by a generalized Fourier series of the form

$$(8) \quad f(x) = c_1 J_n(\lambda_{1n}x) + c_2 J_n(\lambda_{2n}x) + \cdots \quad (n \text{ fixed})$$

it follows from (11) in Sec. 8.12 that the Fourier constants in this so-called Fourier-Bessel series are

$$(9) \quad c_m = \frac{2}{R^2 J_{n+1}^2(\alpha_{mn})} \int_0^R x f(x) J_n(\lambda_{mn}x) dx, \quad m = 1, 2, \dots$$

Here $\lambda_{mn} = \alpha_{mn}/R$.

Problems

1. Graph $J_0(\lambda_{10}x)$, $J_0(\lambda_{20}x)$, $J_0(\lambda_{30}x)$, and $J_0(\lambda_{40}x)$ for $R = 1$ in the interval $0 \leq x \leq 1$. (Use the tables of J_0 and the zeros of J_0 in Secs. 3.5 and 9.10.)

2. Graph $J_1(\alpha_{11}x)$, $J_1(\alpha_{21}x)$, $J_1(\alpha_{31}x)$, and $J_1(\alpha_{41}x)$ in the interval $0 \leq x \leq 1$. Use Table 14 in Sec. 3.5.

Develop the following functions $f(x)$ ($0 < x < R$) in a Fourier-Bessel series of the form

$$f(x) = c_1 J_0(\lambda_{10}x) + c_2 J_0(\lambda_{20}x) + c_3 J_0(\lambda_{30}x) + \cdots$$

and graph the first partial sums.

3. $f(x) = 1$. Hint: use (1), Sec. 3.6. 4. $f(x) = \begin{cases} 1 & \text{if } 0 < x < R/2 \\ 0 & \text{if } R/2 < x < R \end{cases}$

5. $f(x) = \begin{cases} k & \text{if } 0 < x < a \\ 0 & \text{if } a < x < R \end{cases}$ 6. $f(x) = \begin{cases} 0 & \text{if } 0 < x < R/2 \\ k & \text{if } R/2 < x < R \end{cases}$

7. $f(x) = 1 - x^2$ ($R = 1$). Hint: use (1), Sec. 3.6, and integration by parts.

8. $f(x) = R^2 - x^2$ 9. $f(x) = x^2$ 10. $f(x) = x^4$

11. Show that $f(x) = x^n$ ($0 < x < 1$, $n = 0, 1, \dots$) can be represented by the Fourier-Bessel series

$$x^n = \frac{2J_n(\alpha_{1n}x)}{\alpha_{1n}J_{n+1}(\alpha_{1n})} + \frac{2J_n(\alpha_{2n}x)}{\alpha_{2n}J_{n-1}(\alpha_{2n})} + \cdots$$

12. Find a representation of $f(x) = x^n$ ($0 < x < R$, $n = 0, 1, \dots$) similar to that in Prob. 11.

13. It can be shown that the representation in Prob. 11 holds not only for integral $n = 0, 1, \dots$, but for any $n \geq -1/2$. Using (5), Sec. 3.6, show that for $n = -1/2$, after multiplying by \sqrt{x} , the representation becomes the Fourier series

$$\frac{4}{\pi} \left(\cos \frac{\pi}{2}x - \frac{1}{3} \cos \frac{3\pi}{2}x + \frac{1}{5} \cos \frac{5\pi}{2}x - \cdots \right)$$

which represents

$$f(x) = \begin{cases} 1 & \text{if } -1 < x < 1 \\ -1 & \text{if } 1 < x < 3 \end{cases} \quad [f(x+4) = f(x)].$$

14. Represent $f(x) = x^3$ ($0 < x < 2$) by a Fourier-Bessel series involving J_3 .

Chapter 9

Partial Differential Equations

Partial differential equations arise in connection with various physical and geometrical problems when the functions involved depend on two or more independent variables. These variables may be the time and one or several coordinates in space. The present chapter will be devoted to some of the most important partial differential equations occurring in engineering applications. We shall derive these equations from physical principles and consider methods for solving initial and boundary value problems, that is, methods for obtaining solutions of those equations corresponding to the given physical situations.

In Sec. 9.1 we shall define the notion of a solution of a partial differential equation. Sections 9.2–9.4 will be devoted to the one-dimensional wave equation, governing the motion of a vibrating string. The heat equation will be considered in Secs. 9.5 and 9.6, the two-dimensional wave equation (vibrating membranes) in Secs. 9.7–9.10, and Laplace's equation in Secs. 9.11 and 9.12.

Prerequisites for this chapter: ordinary linear differential equations (Chap. 2) and Fourier series (Chap. 8).

Sections which may be omitted in a shorter course: 9.6, 9.9, 9.10.

References: Appendix 1, Part F.

Answers to problems: Appendix 2.

9.1 Basic Concepts

An equation involving one or more partial derivatives of an (unknown) function of two or more independent variables is called a **partial differential equation**. The order of the highest derivative is called the **order of the equation**.

Just as in the case of an ordinary differential equation, we say that a partial differential equation is **linear** if it is of the first degree in the dependent variable and its partial derivatives. If each term of such an equation contains either the dependent variable or one of its derivatives, the equation is said to be **homogeneous**; otherwise it is said to be **nonhomogeneous**.

Example 1. Some important linear partial differential equations of the second order are:

- (1) $\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$ (one-dimensional wave equation)
- (2) $\frac{\partial u}{\partial t} = c^2 \frac{\partial^2 u}{\partial x^2}$ (one-dimensional heat equation)
- (3) $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$ (two-dimensional Laplace equation)
- (4) $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y)$ (two-dimensional Poisson equation)
- (5) $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$ (three-dimensional Laplace equation);

here c is a constant, t is the time, and x, y, z are Cartesian coordinates. Equation (4) (with $f \not\equiv 0$) is nonhomogeneous, while the other equations are homogeneous.

A solution of a partial differential equation in some region R of the space of the independent variables is a function which has all the partial derivatives appearing in the equation and satisfies the equation everywhere in R .

In general, the totality of solutions of a partial differential equation is very large. For example, the functions

$$(6) \quad u = x^2 - y^2, \quad u = e^x \cos y, \quad u = \ln(x^2 + y^2),$$

which are entirely different from each other, are solutions of (3), as the student may verify. We shall see later that the unique solution of a partial differential equation corresponding to a given physical problem will be obtained by the use of additional information arising from the physical situation. For example, in some cases the values of the required solution of the problem on the boundary of some domain will be given ("boundary conditions"); in other cases when the time t is one of the variables, the values of the solution at $t = 0$ will be prescribed ("initial conditions").

We know that if an ordinary differential equation is linear and homogeneous, then from known solutions further solutions can be obtained by superposition. For a homogeneous linear partial differential equation the situation is quite similar. In fact, the following theorem holds.

Fundamental theorem 1. If u_1 and u_2 are any solutions of a linear homogeneous partial differential equation in some region, then

$$u = c_1 u_1 + c_2 u_2,$$

where c_1 and c_2 are any constants, is also a solution of that equation in that region.

The proof of this important theorem is quite similar to that of Theorem 1 in Sec. 2.1 and is left to the student. •

Problems

Verify that the following functions are solutions of the wave equation (1) for a suitable value of the constant c in (1). Sketch figures of these functions as surfaces in space.

1. $u = x^2 + t^2$

2. $u = x^2 + 9t^2$

3. $u = \cos t \sin x$

4. $u = \sin t \sin x$

5. $u = \cos ct \sin x$

6. $u = \sin \omega ct \sin \omega x$

Verify that the following functions are solutions of (1).

7. $u(x, t) = v(x + t) + w(x - t)$

8. $u(x, t) = v(x + ct) + w(x - ct)$

Verify that the following functions are solutions of the heat equation (2).

9. $u = e^{-t} \cos x$

10. $u = e^{-t} \sin x$

11. $u = e^{-\omega^2 c^2 t} \sin \omega x$

Verify that the following functions are solutions of Laplace's equation (3). Sketch figures of the solutions as surfaces in space.

12. $u = x^2 - y^2$

13. $u = x^3 - 3xy^2$

14. $u = 3x^2y - y^3$

15. $u = e^x \cos y$

16. $u = e^x \sin y$

17. $u = \ln(x^2 + y^2)$

18. $u = \arctan(y/x)$

19. $u = \sin x \sinh y$

20. $u = \sin x \cosh y$

21. Verify that $u(x, y) = a \ln(x^2 + y^2) + b$ satisfies Laplace's equation (3) and determine a and b so that u satisfies the boundary conditions $u = 0$ on the circle $x^2 + y^2 = 1$ and $u = 3$ on the circle $x^2 + y^2 = 4$. Sketch a figure of the surface represented by this function.

22. Show that $u = 1/\sqrt{x^2 + y^2 + z^2}$ is a solution of (5).

A partial differential equation involving derivatives with respect to one variable only may be solved like an ordinary differential equation, treating the other independent variables as parameters. Solve the following equations where $u = u(x, y)$.

23. $u_x = 0$

24. $u_y = 0$

25. $u_{xx} = 0$

26. $u_{xx} + u = 0$

Solve the following systems of partial differential equations.

27. (a) $u_{xx} = 0$, (b) $u_{yy} = 0$

28. $u_x = 0$, $u_y = 0$

29. $u_{xx} = 0$, $u_{xy} = 0$, $u_{yy} = 0$

30. $u_{xx} = 0$, $u_{xy} = 0$

Setting $u_x = p$, solve

31. $u_{xy} = u_x$

32. $u_{xy} + u_x = 0$

33. $u_{xy} + u_x + x + y + 1 = 0$

34. Show that if the level curves $z = \text{const}$ of a surface $z = z(x, y)$ are straight lines parallel to the x -axis, then z is a solution of the differential equation $z_x = 0$. Give examples.

35. Show that the solutions $z = z(x, y)$ of $yz_x - xz_y = 0$ represent surfaces of revolution. Give examples. Hint: set $x = r \cos \theta$, $y = r \sin \theta$ and show that the equation becomes $z_\theta = 0$.

9.2 Vibrating String. One-Dimensional Wave Equation

As a first important partial differential equation, let us derive the equation governing small transverse vibrations of an elastic string, which is stretched

to length l and then fixed at the end points. Suppose that the string is distorted and then at a certain instant, say, $t = 0$, it is released and allowed to vibrate. The problem is to determine the vibrations of the string, that is, to find its deflection $u(x, t)$ at any point x and at any time $t > 0$; cf. Fig. 244.

¹¹ When deriving a differential equation corresponding to a given physical problem, we usually have to make simplifying assumptions in order that the resulting equation does not become too complicated. We know this important fact from our considerations of ordinary differential equations, and for partial differential equations the situation is similar.

In our present case we make the following assumptions.

1. *The mass of the string per unit length is constant ("homogeneous string"). The string is perfectly elastic and does not offer any resistance to bending.*
2. *The tension caused by stretching the string before fixing it at the end points is so large that the action of the gravitational force on the string can be neglected.*
3. *The motion of the string is a small transverse vibration in a vertical plane, that is, each particle of the string moves strictly vertically, and the deflection and the slope at any point of the string are small in absolute value.*

These assumptions are such that we may expect that the solution $u(x, t)$ of the differential equation to be obtained will reasonably well describe small vibrations of the physical "nonidealized" string of small homogeneous mass under large tension.

To obtain the differential equation we consider the forces acting on a small portion of the string (Fig. 244). Since the string does not offer resistance to bending, the tension is tangential to the curve of the string at each point. Let T_1 and T_2 be the tensions at the end points P and Q of that portion. Since there is no motion in horizontal direction, the horizontal components of the tension must be constant. Using the notation shown in Fig. 244, we thus obtain

$$(1) \quad T_1 \cos \alpha = T_2 \cos \beta = T = \text{const.}$$

In vertical direction we have two forces, namely the vertical components $-T_1 \sin \alpha$ and $T_2 \sin \beta$ of T_1 and T_2 ; here the minus sign appears because that component at P is directed downward. By Newton's second law the

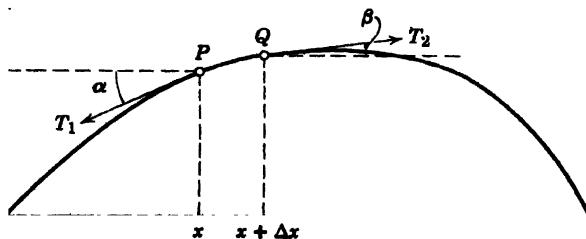


Fig. 244. Vibrating string.

resultant of those two forces is equal to the mass $\rho \Delta x$ of the portion times the acceleration $\partial^2 u / \partial t^2$, evaluated at some point between x and $x + \Delta x$; here ρ is the mass of the undeflected string per unit length, and Δx is the length of the portion of the undeflected string. Hence

$$T_2 \sin \beta - T_1 \sin \alpha = \rho \Delta x \frac{\partial^2 u}{\partial t^2}$$

By using (1) we obtain

$$\frac{T_2 \sin \beta}{T_2 \cos \beta} - \frac{T_1 \sin \alpha}{T_1 \cos \alpha} = \tan \beta - \tan \alpha = \frac{\rho \Delta x}{T} \frac{\partial^2 u}{\partial t^2}$$

Now $\tan \alpha$ and $\tan \beta$ are the slopes of the curve of the string at x and $x + \Delta x$, that is,

$$\tan \alpha = \left(\frac{\partial u}{\partial x} \right)_x \quad \text{and} \quad \tan \beta = \left(\frac{\partial u}{\partial x} \right)_{x+\Delta x}$$

Here we have to write *partial* derivatives because u depends also on t . Dividing (2) by Δx , we thus obtain

$$\frac{1}{\Delta x} \left[\left(\frac{\partial u}{\partial x} \right)_{x+\Delta x} - \left(\frac{\partial u}{\partial x} \right)_x \right] = \frac{\rho}{T} \frac{\partial^2 u}{\partial t^2}.$$

If we let Δx approach zero, we obtain the linear homogeneous partial differential equation

$$(3) \quad \frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \quad c^2 = \frac{T}{\rho}$$

This is the so-called **one-dimensional wave equation**, which governs our problem. The notation c^2 (instead of c) for the physical constant T/ρ has been chosen to indicate that this constant is positive.

Solutions of the equation will be obtained in the following section.

9.3 Separation of Variables (Product Method)

We have seen that the vibrations of an elastic string are governed by the one-dimensional wave equation

$$(1) \quad \frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2},$$

where $u(x, t)$ is the deflection of the string. Since the string is fixed at the ends $x = 0$ and $x = l$, we have the two **boundary conditions**

$$(2) \quad u(0, t) = 0, \quad u(l, t) = 0 \quad \text{for all } t.$$

The form of the motion of the string will depend on the initial deflection (deflection at $t = 0$) and on the initial velocity (velocity at $t = 0$). Denoting the initial deflection by $f(x)$ and the initial velocity by $g(x)$, we thus obtain the two initial conditions

$$(3) \quad u(x, 0) = f(x)$$

and

$$(4) \quad \frac{\partial u}{\partial t} \Big|_{t=0} = g(x).$$

Our problem is now to find a solution of (1) satisfying the conditions (2)–(4). We shall proceed step by step, as follows.

First step. By applying the so-called *product method*, or *method of separating variables*, we shall obtain two ordinary differential equations.

Second step. We shall determine solutions of those equations that satisfy the boundary conditions.

Third step. Those solutions will be composed so that the result will be a solution of the wave equation (1), satisfying also the given initial conditions.

Let us start in detail with the first step.

First step. The product method yields solutions of (1) of the form

$$(5) \quad u(x, t) = F(x)G(t)$$

which are a product of two functions, each depending only on one of the variables x and t . We shall see later that this method has various applications in engineering mathematics. By differentiating (5) we obtain

$$\frac{\partial^2 u}{\partial t^2} = FG \quad \text{and} \quad \frac{\partial^2 u}{\partial x^2} = F''G,$$

where dots denote derivatives with respect to t and primes derivatives with respect to x . By inserting this into (1) we have

$$F\ddot{G} = c^2 F''G.$$

Dividing by c^2FG , we find

$$\frac{\ddot{G}}{c^2 G} = \frac{F''}{F}.$$

The expression on the left involves functions depending only on t while the expression on the right involves functions depending only on x . Hence both expressions must be equal to a constant, say, k , because if the expression on the left is not constant, then changing t will presumably change the value of this expression but certainly not that on the right, since the latter does not depend on t . Similarly, if the expression on the right is not constant, changing x will presumably change the value of this expression but certainly not that on the left. Thus

$$\frac{\ddot{G}}{c^2 G} = \frac{F''}{F} = k.$$

This yields immediately the two ordinary linear differential equations,

$$(6) \quad F'' - kF = 0$$

and

$$(7) \quad \ddot{G} - c^2kG = 0.$$

In these equations, k is still arbitrary.

Second step. We shall now determine solutions F and G of (6) and (7) so that $u = FG$ satisfies (2), that is,

$$u(0, t) = F(0)G(t) = 0, \quad u(l, t) = F(l)G(t) = 0 \quad \text{for all } t$$

Clearly, if $G \equiv 0$, then $u \equiv 0$, which is of no interest. Thus $G \not\equiv 0$ and then

$$(8) \quad (a) \quad F(0) = 0 \quad (b) \quad F(l) = 0.$$

For $k = 0$ the general solution of (6) is $F = ax + b$, and from (8) we obtain $a = b = 0$. Hence $F \equiv 0$, which is of no interest because then $u \equiv 0$. For positive $k = \mu^2$ the general solution of (6) is

$$F = Ae^{\mu x} + Be^{-\mu x},$$

and from (8) we obtain $F \equiv 0$, as before. Hence we are left with the possibility of choosing k negative, say, $k = -p^2$. Then (6) takes the form

$$F'' + p^2F = 0,$$

and the general solution is

$$F(x) = A \cos px + B \sin px.$$

From this and (8) we have

$$F(0) = A = 0 \quad \text{and then} \quad F(l) = B \sin pl = 0.$$

We must take $B \neq 0$ since otherwise $F \equiv 0$. Hence $\sin pl = 0$, that is,

$$(9) \quad pl = n\pi \quad \text{or} \quad p = \frac{n\pi}{l} \quad (\text{n integral}).$$

Setting $B = 1$, we thus obtain infinitely many solutions $F(x) = F_n(x)$,

$$(10) \quad F_n(x) = \sin \frac{n\pi}{l} x \quad n = 1, 2, \dots,$$

which satisfy (8). (For negative integral n we obtain essentially the same solutions, except for a minus sign, because $\sin(-x) = -\sin x$.)

k is now restricted to the values $k = -p^2 = -(n\pi/l)^2$, resulting from (9). For these k the equation (7) takes the form

$$\ddot{G} + \lambda_n^2 G = 0 \quad \text{where} \quad \lambda_n = \frac{cn\pi}{l}.$$

The general solution is

$$G_n(t) = B_n \cos \lambda_n t + B_n^* \sin \lambda_n t.$$



Fig. 245. Normal modes of the vibrating string.

Hence the functions $u_n(x, t) = F_n(x)G_n(t)$, at length

$$(11) \quad u_n(x, t) = (B_n \cos \lambda_n t + B_n^* \sin \lambda_n t) \sin \frac{n\pi}{l} x \quad (n = 1, 2, \dots)$$

are solutions of (1), satisfying the boundary conditions (2). These functions are called the **eigenfunctions**, or *characteristic functions*, and the values $\lambda_n = cn\pi/l$ are called the **eigenvalues**, or *characteristic values*, of the vibrating string. The set $\lambda_1, \lambda_2, \dots$ is called the *spectrum*.

We see that each u_n represents a harmonic motion having the frequency $\lambda_n/2\pi = cn/2l$ cycles per unit time. The motion is called the *n*th normal mode of the string. The first normal mode is known as the *fundamental mode* ($n = 1$), and the others as *overtones*; musically they give the octave, octave plus fifth, etc. Since in (11)

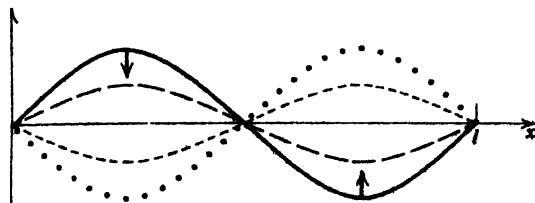
$$\sin \frac{n\pi x}{l} = 0 \quad \text{at} \quad x = \frac{l}{n}, \frac{2l}{n}, \dots, \frac{n-1}{n}l,$$

the *n*th normal mode has $n - 1$ so-called **nodes**, that is, points of the string which do not move (Fig. 245).

Figure 246 shows the second normal mode for various values of t . At any instant the string has the form of a sine wave. When the left part of the string is moving downward the other half is moving upward, and conversely. For the other modes the situation is similar.

Third step. Clearly, a single solution $u_n(x, t)$ will, in general, not satisfy the initial conditions (3) and (4). Now, since the equation (1) is linear and homogeneous, it follows from Fundamental Theorem 1 in Sec. 9.1 that the sum of finitely many solutions u_n is a solution of (1). To obtain a solution that satisfies (3) and (4), we consider the infinite series

$$(12) \quad u(x, t) = \sum_{n=1}^{\infty} u_n(x, t) = \sum_{n=1}^{\infty} (B_n \cos \lambda_n t + B_n^* \sin \lambda_n t) \sin \frac{n\pi}{l} x.$$

Fig. 246. Second normal mode for various values of t .

From this and (3) it follows that

$$(13) \quad u(x, 0) = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi}{l} x = f(x).$$

Hence, in order that (12) satisfy (3), the coefficients B_n must be chosen so that $u(x, 0)$ becomes a half-range expansion of $f(x)$, namely, the Fourier sine series of $f(x)$; that is [cf. (4) in Sec. 8.5]

$$(14) \quad B_n = \frac{2}{l} \int_0^l f(x) \sin \frac{n\pi x}{l} dx, \quad n = 1, 2, \dots$$

Similarly, by differentiating (12) with respect to t and using (4) we find

$$\begin{aligned} \frac{\partial u}{\partial t} \Big|_{t=0} &= \left[\sum_{n=1}^{\infty} (-B_n \lambda_n \sin \lambda_n t + B_n * \lambda_n \cos \lambda_n t) \sin \frac{n\pi x}{l} \right]_{t=0} \\ &= \sum_{n=1}^{\infty} B_n * \lambda_n \sin \frac{n\pi x}{l} = g(x). \end{aligned}$$

Hence, in order that (12) satisfy (4), the coefficients $B_n *$ must be chosen so that, for $t = 0$, $\partial u / \partial t$ becomes the Fourier sine series of $g(x)$; thus, by (4) in Sec. 8.5,

$$B_n * \lambda_n = \frac{2}{l} \int_0^l g(x) \sin \frac{n\pi x}{l} dx$$

or, since $\lambda_n = cn\pi/l$,

$$(15) \quad B_n * = \frac{2}{cn\pi} \int_0^l g(x) \sin \frac{n\pi x}{l} dx, \quad n = 1, 2, \dots$$

It follows that $u(x, t)$, given by (12) with coefficients (14) and (15), is a solution of (1) that satisfies the conditions (2)–(4), provided that the series (12) converges and also that the series obtained by differentiating (12) twice (termwise) with respect to x and t , converge and have the sums $\partial^2 u / \partial x^2$ and $\partial^2 u / \partial t^2$, respectively, which are continuous.

Hence the solution (12) is at first a purely formal expression, and we shall now establish it. For the sake of simplicity we consider only the case when the initial velocity $g(x)$ is identically zero. Then the $B_n *$ are zero, and (12) reduces to the form

$$(16) \quad u(x, t) = \sum_{n=1}^{\infty} B_n \cos \lambda_n t \sin \frac{n\pi x}{l}, \quad \lambda_n = \frac{cn\pi}{l}.$$

It is possible to sum this series, that is, to write the result in a closed or finite form. From (11b) in Sec. 0.1 it follows that

$$\cos \frac{cn\pi}{l} t \sin \frac{n\pi}{l} x = \frac{1}{2} \left[\sin \left\{ \frac{n\pi}{l} (x - ct) \right\} + \sin \left\{ \frac{n\pi}{l} (x + ct) \right\} \right].$$

Fig. 247. Odd periodic extension of $f(x)$.

Consequently we may write (16) in the form

$$u(x, t) = \frac{1}{2} \sum_{n=1}^{\infty} B_n \sin \left\{ \frac{n\pi}{l} (x - ct) \right\} + \frac{1}{2} \sum_{n=1}^{\infty} B_n \sin \left\{ \frac{n\pi}{l} (x + ct) \right\}.$$

These two series are those obtained by substituting $x - ct$ and $x + ct$, respectively, for the variable x in the Fourier sine series (13) for $f(x)$. Therefore,

$$(17) \quad u(x, t) = \frac{1}{2}[f^*(x - ct) + f^*(x + ct)],$$

where f^* is the odd periodic extension of f with the period $2l$ (Fig. 247). Since the initial deflection $f(x)$ is continuous on the interval $0 \leq x \leq l$ and zero at the end points, it follows from (17) that $u(x, t)$ is a continuous function of both variables x and t for all values of the variables. By differentiating (17) we see that $u(x, t)$ is a solution of (1), provided $f(x)$ is twice differentiable on the interval $0 < x < l$, and has one-sided second derivatives at $x = 0$ and $x = l$, which are zero. Under these conditions $u(x, t)$ is established as a solution of (1), satisfying (2)–(4).

If $f'(x)$ and $f''(x)$ are merely piecewise continuous (cf. Sec. 4.1), or if those one-sided derivatives are not zero, then for each t there will be finitely many values of x at which the second derivatives of u appearing in (1) do not exist. Except at these points the wave equation will still be satisfied, and we may then regard $u(x, t)$ as a solution of our problem in a broader sense. For example, the case of a triangular initial deflection (Ex. 1, below) leads to a solution of this type.

Let us mention a very interesting physical interpretation of (17). The graph of $f^*(x - ct)$ is obtained from the graph of $f^*(x)$ by shifting the latter ct units to the right (Fig. 248). This means that $f^*(x - ct)$ ($c > 0$) represents a wave which is traveling to the right as t increases. Similarly, $f^*(x + ct)$ represents a wave which is traveling to the left, and $u(x, t)$ is the superposition of these two waves.

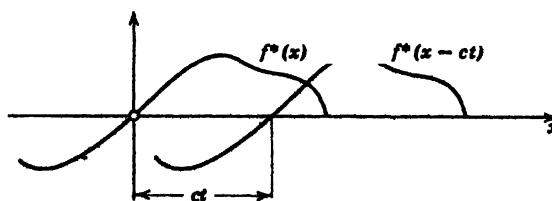


Fig. 248. Interpretation of (17).

Example 1. Find the solution of the wave equation (1) corresponding to the triangular initial deflection

$$f(x) = \begin{cases} \frac{2k}{l}x & \text{when } 0 < x < \frac{l}{2} \\ \frac{2k}{l}(l-x) & \text{when } \frac{l}{2} < x < l \end{cases}$$

and initial velocity zero. Since $g(x) \equiv 0$, we have $B_n^* = 0$ in (12), and from Ex. 1 in Sec. 8.5 we see that the B_n are given by (5), Sec. 8.5. Thus (12) takes the form

$$u(x, t) = \frac{8k}{\pi^2} \left[\frac{1}{1^2} \sin \frac{\pi}{l} x \cos \frac{\pi c}{l} t - \frac{1}{3^2} \sin \frac{3\pi}{l} x \cos \frac{3\pi c}{l} t + \dots \right].$$

For plotting the graph of the solution we may use $u(x, 0) = f(x)$ and the above interpretation of the two functions in the representation (17). This leads to the graph shown in Fig. 249.

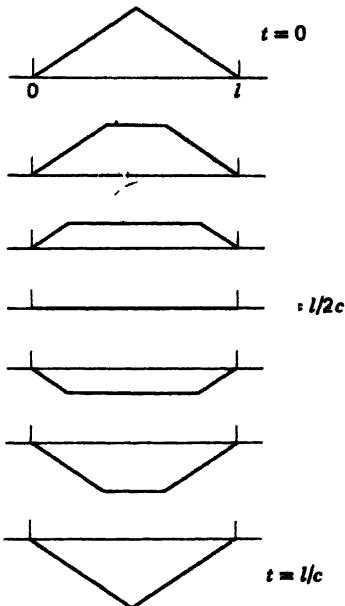


Fig. 249. Solution in Example 1 for various values of t .

Problems

- Using the interpretation of the functions in (17), graph the solution in Ex. 1 and show that this leads to Fig. 249.
- In what manner does the frequency of the fundamental mode of the vibrating string depend on the length of the string, the tension, and the mass of the string per unit length?

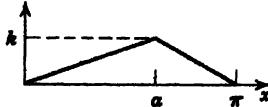
Find the deflection $u(x, t)$ of the vibrating string (length $l = \pi$, ends fixed, and $c^2 = T/\rho = 1$) corresponding to zero initial velocity and initial deflection:

3. $0.01 \sin x$

4. $k \sin 2x$

5. $k(\sin x + \sin 3x)$

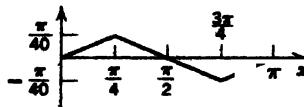
6.



8. $0.01x(\pi - x)$

9. $0.01x(\pi^2 - x^2)$

7.



10. $k[(\frac{1}{2}\pi)^4 - (x - \frac{1}{2}\pi)^4]$

- What is the ratio of the amplitudes of the fundamental mode and the second overtone in Prob. 8? The ratio $a_1^2/(a_1^2 + a_2^2 + \dots)$? Hint: use (8) in Sec. 8.8, with the equality sign.

Separating variables, find solutions $u(x, y)$ of the following equations.

12. $u_x + u_y = 0$

13. $u_x = u_y$

14. $u_x = yu_y$

15. $xu_x = yu_y$

16. $ayu_x = buu_y$

17. $u_x + u_y = 2(x + y)u$

18. $u_x = u$

19. $u_{xx} + u_{yy} = 0$

20. $x^2u_{xy} + 3y^2u = 0$

21. Show that substitution of

$$(18) \quad u(x, t) = \sum_{n=0}^{\infty} G_n(t) \sin \frac{n\pi x}{l}$$

into the wave equation (1) leads to the equation

$$\ddot{G}_n + \lambda_n^2 G_n = 0, \quad \lambda_n = \frac{cn\pi}{l}.$$

22. (Forced vibrations of an elastic string) Show that the forced vibrations of an elastic string under an external force $P(x, t)$ per unit length acting normal to the string are governed by the equation

$$u_{tt} = c^2 u_{xx} + \frac{P}{\rho}.$$

Consider and solve the equation in Prob. 22 for the sinusoidal force $P = A\rho \sin \omega t$, as follows.

23. Show that

$$P/\rho = A \sin \omega t = \sum_{n=1}^{\infty} k_n(t) \sin \frac{n\pi x}{l}$$

where $k_n(t) = (2A/n\pi)(1 - \cos n\pi) \sin \omega t$; consequently $k_n = 0$ (n even), and $k_n = (4A/n\pi) \sin \omega t$ (n odd).

24. Show that by substituting the expressions for u in Prob. 21 and P/ρ in Prob. 23 into the equation under consideration we obtain

$$\ddot{G}_n + \lambda_n^2 G_n = \frac{2A}{n\pi} (1 - \cos n\pi) \sin \omega t, \quad \lambda_n = \frac{cn\pi}{l}.$$

Show that if $\lambda_n^2 \neq \omega^2$, the solution is

$$G_n(t) = B_n \cos \lambda_n t + B_n^* \sin \lambda_n t + \frac{2A(1 - \cos n\pi)}{n\pi(\lambda_n^2 - \omega^2)} \sin \omega t.$$

25. Determine B_n and B_n^* in Prob. 24 so that u satisfies the initial conditions $u(x, 0) = f(x)$, $u_t(x, 0) = 0$.

26. Show that in the case of resonance ($\lambda_n = \omega$),

$$G_n(t) = B_n \cos \omega t + B_n^* \sin \omega t - \frac{A}{n\pi\omega} (1 - \cos n\pi) t \cos \omega t.$$

27. (Vibrations of a beam) It can be shown that the small free vertical vibrations of a uniform cantilever beam are governed by the fourth-order equation

$$(19) \quad \frac{\partial^2 u}{\partial t^2} + c^2 \frac{\partial^4 u}{\partial x^4} = 0 \quad (\text{Ref. [F8]}),$$

where $c^2 = EI/\rho A$ (E = Young's modulus of elasticity, I = moment of inertia of the cross section with respect to the y -axis in Fig. 250 (cf. p. 498), ρ = density, A = cross-sectional area). Substituting $u = F(x)G(t)$ into (19) and separating variables, show that

$$F^{(4)}/F = -\ddot{G}/c^2 G = \beta^4 = \text{const},$$

$$F(x) = A \cos \beta x + B \sin \beta x + C \cosh \beta x + D \sinh \beta x,$$

$$G(t) = a \cos c\beta^2 t + b \sin c\beta^2 t.$$

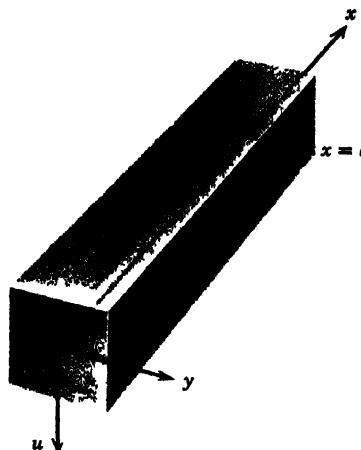


Fig. 250. Undeformed beam.

28. Find solutions $u_n = F_n(x)G_n(t)$ of (19) corresponding to zero initial velocity and satisfying the boundary conditions (Fig. 251)
 $u(0, t) = 0, u(l, t) = 0$ (end fixed for all times t),
 $u_{xx}(0, t) = 0, u_{xx}(l, t) = 0$ (zero moments and, hence, zero curvature at the ends).

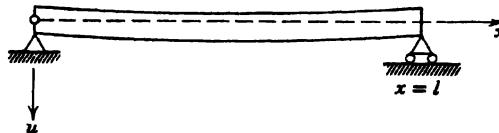


Fig. 251. Beam in Problem 28.

29. Find the solution of (19) that satisfies the conditions in Prob. 28 and the initial condition $u(x, 0) = f(x) = x(l - x)$.
 30. Compare the results of Probs. 29 and 8. What is the basic difference between the frequencies of the normal modes of the vibrating string and the vibrating beam?
 31. What are the boundary conditions if the beam is clamped at both ends (Fig. 252)?

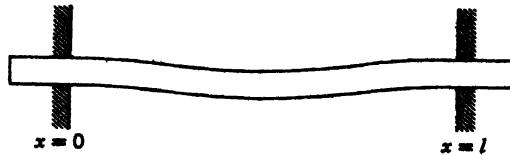


Fig. 252. Beam clamped at both ends.

32. Show that $F(x)$ in Prob. 27 satisfies the conditions in Prob. 31, if βl is a root of the equation

$$\cosh \beta l / \cos \beta l = 1.$$

 33. Determine approximate solutions of (20).

34. If the beam is clamped at the left and free at the other end (Fig. 253), the boundary conditions are

$$u(0, t) = 0, \quad u_x(0, t) = 0, \quad u_{xx}(l, t) = 0, \quad u_{xxx}(l, t) = 0.$$

Show that $F(x)$ in Prob. 27 satisfies these conditions, if βl is a root of the equation

$$\cosh \beta l / \cos \beta l = -1.$$



Fig. 253. Beam clamped at one end and free at the other.

35. Find approximate solutions of the equation in Prob. 34.

✓ 9.4 D'Alembert's Solution of the Wave Equation

It is interesting to note that the solution (17) in the last section of the wave equation

$$(1) \quad \frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \quad c^2 = \frac{T}{\rho}$$

can be immediately obtained by transforming (1) in a suitable way, namely, by introducing the new independent variables¹

$$(2) \quad v = x + ct, \quad z = x - ct.$$

u then becomes a function of v and z , and the derivatives in (1) can be expressed in terms of derivatives with respect to v and z by the use of the chain rule in Sec. 5.15. Denoting partial derivatives by subscripts, we see from (2) that $v_x = 1$ and $z_x = 1$ and, therefore,

$$u_x = u_v v_x + u_z z_x = u_v + u_z.$$

By applying the chain rule to the right side we find

$$u_{xx} = (u_v + u_z)_x = (u_v + u_z)_v v_x + (u_v + u_z)_z z_x.$$

Since $v_x = 1$ and $z_x = 1$, this becomes

$$u_{xx} = u_{vv} + 2u_{vz} + u_{zz}.$$

The other derivative in (1) is transformed by the same procedure, and the result is

$$u_{tt} = c^2(u_{vv} - 2u_{vz} + u_{zz}).$$

¹ We mention that the general theory of partial differential equations provides a systematic way for finding this transformation which will simplify the equation. Cf. Ref. [F7] in Appendix 1.

By inserting these two results in (1) we obtain

$$(3) \quad u_{vv} = \frac{\partial^2 u}{\partial v \partial z} = 0.$$

We may integrate this equation with respect to z , finding

$$\frac{\partial u}{\partial v} = h(v)$$

where $h(v)$ is an arbitrary function of v . Integrating this with respect to v , we have

$$u = \int h(v) dv + \psi(z)$$

where $\psi(z)$ is an arbitrary function of z . Since the integral is a function of v , say, $\phi(v)$, the solution u is of the form

$$u = \phi(v) + \psi(z).$$

Because of (2) we may write

$$(4) \quad u(x, t) = \phi(x + ct) + \psi(x - ct).$$

This is known as d'Alembert's solution² of the wave equation (1).

The functions ϕ and ψ can be determined from the initial conditions. Let us illustrate this in the case of zero initial velocity and given initial deflection $u(x, 0) = f(x)$.

By differentiating (4) we have

$$(5) \quad \frac{\partial u}{\partial t} = c\phi'(x + ct) - c\psi'(x - ct)$$

where primes denote derivatives with respect to the *entire* arguments $x + ct$ and $x - ct$, respectively. From (4), (5), and the initial conditions we have

$$u(x, 0) = \phi(x) + \psi(x) = f(x)$$

$$u_t(x, 0) = c\phi'(x) - c\psi'(x) = 0.$$

From the last equation, $\psi' = \phi'$. Hence $\psi = \phi + k$, and from this and the first equation, $2\phi + k = f$ or $\phi = (f - k)/2$. With these functions ϕ and ψ the solution (4) becomes

$$(6) \quad u(x, t) = \frac{1}{2}[f(x + ct) + f(x - ct)],$$

in agreement with (17) in Sec. 9.3. The student may show that because of the boundary conditions (2) in that section the function f must be odd and have period $2l$.

Our result shows that the two initial conditions and the boundary conditions determine the solution uniquely.

² JEAN-LE-ROND D'ALEMBERT (1717-1783), French mathematician, who is known for his important work in mechanics.

Problems

1. Express x and t in terms of v and z , cf. (2), and use the result for transforming (3) into (1).

Using (6), sketch a figure (of the type of Fig. 249 in Sec. 9.3) of the deflection $u(x, t)$ of a vibrating string (length $l = 1$, ends fixed) starting with initial velocity zero and initial deflection $f(x)$, where

$$\begin{array}{lll} 2. f(x) = k \sin 2\pi x & 3. f(x) = kx(1-x) & 4. f(x) = k(x-x^3) \\ 5. f(x) = k(x^2 - x^4) & 6. f(x) = k(x^3 - x^5) & 7. f(x) = k \sin^2 \pi x \end{array}$$

Using the indicated transformations, solve the following equations.

$$\begin{array}{lll} 8. xu_{xy} = yu_{yy} + u_y & (v = x, z = xy) \\ 9. u_{xy} - u_{yy} = 0 & (v = x, z = x + y) \\ 10. u_{xx} + 2u_{xy} + u_{yy} = 0 & (v = x, z = x - y) \\ 11. u_{xx} - 2u_{xy} + u_{yy} = 0 & (v = x, z = x + y) \\ 12. u_{xx} + u_{xy} - 2u_{yy} = 0 & (v = x + y, z = 2x - y) \end{array}$$

13. (Types of linear partial differential equations) An equation of the form

$$(7) \quad Au_{xx} + 2Bu_{xy} + Cu_{yy} = F(x, y, u, u_x, u_y)$$

is said to be **elliptic** if $AC - B^2 > 0$, **parabolic** if $AC - B^2 = 0$, and **hyperbolic** if $AC - B^2 < 0$. (Here A, B, C may be functions of x and y , and the type of (7) may be different in different parts of the xy -plane.) Find whether the equations in Probs. 8-12 are elliptic, parabolic, or hyperbolic.

14. If (7) is *hyperbolic*, it can be transformed to the *normal form*

$$u_{vv} = F^*(v, z, u, u_v, u_z)$$

by setting $v = \Phi(x, y)$, $z = \Psi(x, y)$ where $\Phi = \text{const}$ and $\Psi = \text{const}$ are the solutions $y = y(x)$ of $Ay'^2 - 2By' + C = 0$ (cf. Ref. [F 7]). Show that in the case of the wave equation (1),

$$\Phi = x + ct, \quad \Psi = x - ct.$$

15. If (7) is *parabolic*, the substitution $v = x$, $z = \Psi(x, y)$, with Ψ defined as in Prob. 14, reduces it to the *normal form* $u_{vv} = F^*(v, z, u, u_v, u_z)$. Verify this result for the equation in Prob. 10.

✓9.5 One-Dimensional Heat Flow

The heat flow in a body of homogeneous material is governed by the heat equation (cf. Sec. 6.9)

$$\frac{\partial u}{\partial t} = c^2 \nabla^2 u \quad c^2 = \frac{K}{\sigma\rho}$$

where $u(x, y, z, t)$ is the temperature in the body, K is the thermal conductivity, σ is the specific heat, and ρ is the density of material of the body. $\nabla^2 u$ is the Laplacian of u , and with respect to Cartesian coordinates x, y, z ,

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}.$$

$$\nabla^2 u = \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2}$$



Fig. 254. Bar under consideration.

As an important application, let us consider the temperature in a long thin bar or wire of constant cross section and homogeneous material which is oriented along the x -axis (Fig. 254) and is perfectly insulated laterally, so that heat flows in the x -direction only. Then u depends only on x and time t , and the heat equation becomes the so-called **one-dimensional heat equation**

$$(1) \quad \frac{\partial u}{\partial t} = c^2 \frac{\partial^2 u}{\partial x^2}.$$

While the wave equation (cf. Sec. 9.2) involves the *second* partial derivative $\partial^2 u / \partial t^2$, the heat equation involves the *first* derivative $\partial u / \partial t$, and we shall see that the solutions of (1) are entirely different from those of the wave equation, although the procedure of solving (1) is quite similar to that in the case of the wave equation. We shall solve (1) for some important types of boundary and initial conditions.

Let us start with the case when the ends $x = 0$ and $x = l$ of the bar are kept at temperature zero. Then the *boundary conditions* are

$$(2) \quad u(0, t) = 0, \quad u(l, t) = 0 \quad \text{for all } t.$$

Let $f(x)$ be the initial temperature in the bar. Then the *initial condition* is

$$(3) \quad u(x, 0) = f(x),$$

where $f(x)$ is a given function. We shall determine a solution $u(x, t)$ of (1), satisfying (2) and (3).

First step. Using the method of separating variables, we first determine solutions of (1) that satisfy the boundary conditions (2). We start from

$$(4) \quad u(x, t) = F(x)G(t).$$

Differentiating and substituting this into (1), we obtain

$$\dot{F}\dot{G} = c^2 F''G$$

where dots denote derivatives with respect to t and primes denote derivatives with respect to x . Dividing by $c^2 FG$, we have

$$(5) \quad \frac{\dot{G}}{c^2 G} = \frac{F''}{F}.$$

The expression on the left depends only on t , while the right side depends only on x . As in Sec. 9.3, we conclude that both expressions must be equal to a constant, say, k . The student may show that for $k \geq 0$ the only solution

$u = FG$ that satisfies (2) is $u \equiv 0$. For negative $k = -p^2$ we obtain from (5)

$$\frac{\dot{G}}{c^2 G} = \frac{F''}{F} = -p^2$$

and from this the two ordinary differential equations

$$(6) \quad F'' + p^2 F = 0$$

and

$$(7) \quad \dot{G} + c^2 p^2 G = 0.$$

Second step. We consider (6). The general solution is

$$(8) \quad F(x) = A \cos px + B \sin px.$$

From (2) it follows that

$$u(0, t) = F(0)G(t) = 0 \quad \text{and} \quad u(l, t) = F(l)G(t) = 0.$$

Since $G \equiv 0$ implies $u \equiv 0$, we require that $F(0) = 0$ and $F(l) = 0$. By (8), $F(0) = A$. Thus $A = 0$, and, therefore,

$$F(l) = B \sin pl.$$

We must have $B \neq 0$, since otherwise $F \equiv 0$. Hence the condition $F(l) = 0$ leads to

$$\sin pl = 0 \quad \text{or} \quad p = \frac{n\pi}{l}, \quad n = 1, 2, \dots$$

Setting $B = 1$, we thus obtain the solutions

$$F_n(x) = \sin \frac{n\pi x}{l} \quad n = 1, 2, \dots$$

of (6), satisfying (2). (As in Sec. 9.3, we need not consider negative integral values of n .)

For the values $p = n\pi/l$ the equation (7) takes the form

$$\dot{G} + \lambda_n^2 G = 0 \quad \text{where} \quad \lambda_n = \frac{cn\pi}{l}.$$

The general solution is

$$G_n(t) = B_n e^{-\lambda_n^2 t}, \quad n = 1, 2, \dots$$

where B_n is a constant. Hence the functions

$$(9) \quad u_n(x, t) = F_n(x) G_n(t) = B_n \sin \frac{n\pi x}{l} e^{-\lambda_n^2 t} \quad n = 1, 2, \dots$$

are solutions of the heat equation (1), satisfying (2).

Third step. To find a solution also satisfying (3), we consider the series

$$(10) \quad u(x, t) = \sum_{n=1}^{\infty} u_n(x, t) = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{l} e^{-\lambda_n^2 t} \quad \left(\lambda_n = \frac{cn}{l} \right)$$

From this and (3) it follows that

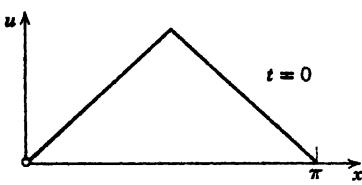
$$u(x, 0) = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{l} = f(x).$$

Hence, for (10) to satisfy (3), the coefficients B_n must be chosen such that $u(x, 0)$ becomes a half-range expansion of $f(x)$, namely, the Fourier sine series of $f(x)$; that is [cf. (4) in Sec. 8.5],

$$(11) \quad B_n = \frac{2}{l} \int_0^l f(x) \sin \frac{n\pi x}{l} dx \quad n = 1, 2, \dots$$

The solution of our problem can be established, assuming that $f(x)$ is piecewise continuous on the interval $0 \leq x \leq l$ (cf. p. 195), and has one-sided derivatives³ at all interior points of that interval; that is, under the assumptions the series (10) with coefficients (11) is the solution of a physical problem. The proof, which requires the knowledge of uniform convergence of series, will be given at a later occasion (Probs. 19, 20 at the end of Sec. 14.4).

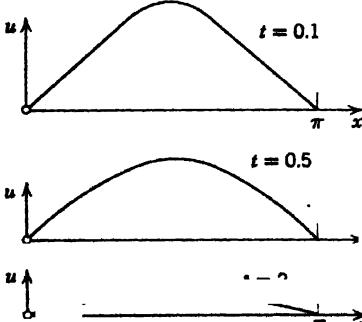
Because of the exponential factor all the terms in (11) approach zero as t approaches infinity. The rate of decay varies with n .



Example 1. If the initial temperature is

$$f(x) = \begin{cases} x & \text{when } 0 < x < l/2, \\ l-x & \text{when } l/2 < x < l \end{cases}$$

(cf. Fig. 255 where $l = \pi$ and $c = 1$), then obtain from (11)



$$(12) \quad B_n = \frac{2}{l} \left(\int_0^{l/2} x \sin \frac{n\pi x}{l} dx + \int_{l/2}^l (l-x) \sin \frac{n\pi x}{l} dx \right).$$

Integration yields $B_n = 0$ when n is even and

$$B_n = \frac{4l}{n^2 \pi^2} \quad (n = 1, 3, 5, \dots)$$

$$B_n = -\frac{4l}{n^2 \pi^2} \quad (n = 2, 4, 6, \dots)$$

Fig. 255. Solution of Example 1 for various values of t .

³ Cf. footnote 4 in Sec. 8.2.

Hence the solution is

$$u(x, t) = \frac{4l}{\pi^2} \left[\sin \frac{\pi x}{l} e^{-(c\pi/l)^2 t} - \frac{1}{9} \sin \frac{3\pi x}{l} e^{-(3c\pi/l)^2 t} + \dots \right].$$

The reader may compare Fig. 255 and Fig. 249 in Sec. 9.3.

Problems

1. Show that the one-dimensional heat equation is parabolic and the one-dimensional wave equation is hyperbolic (cf. Prob. 13 at the end of the last section). What is the essential difference between the solutions of these two equations?
2. How does the rate of decay of (9) for fixed n depend on the specific heat, the density, and the thermal conductivity of the material?
3. Graph u_1, u_2, u_3 [cf. (9), with $B_n = 1, c = 1, l = \pi$] as functions of x for the values $t = 0, 1, 2, 3$. Compare the behavior of these functions.
4. Represent the solutions in Prob. 3 graphically as surfaces over the xt -plane.
5. The ends of a rod that satisfies the assumptions in the text are kept at different constant temperatures $u(0, t) = U_1$ and $u(l, t) = U_2$. Find the temperature $u_l(x)$ in the rod after a long time (theoretically: as $t \rightarrow \infty$).

Find the temperature $u(x, t)$ in a bar of silver (length 10 cm, constant cross section of area 1 cm², density 10.6 gm/cm³, thermal conductivity 1.04 cal/cm deg sec, specific heat 0.056 cal/gm deg) which is perfectly insulated laterally, whose ends are kept at temperature 0°C and whose initial temperature (in °C) is $f(x)$, where

6. $f(x) = \sin 0.2\pi x$
7. $f(x) = \sin 0.1\pi x$
8. $f(x) = \begin{cases} x & \text{if } 0 < x < 5 \\ 0 & \text{if } 5 < x < 10 \end{cases}$
9. $f(x) = \begin{cases} x & \text{if } 0 < x < 5 \\ 10 - x & \text{if } 5 < x < 10 \end{cases}$
10. $f(x) = x(100 - x^2)$
11. $f(x) = x(10 - x)$

12. Find the temperature $u(x, t)$ in a bar of length l which is perfectly insulated, also at the ends at $x = 0$ and $x = l$, assuming that $u(x, 0) = f(x)$. Physical information: the flux of heat through the faces at the ends is proportional to the values of $\partial u / \partial x$ there. Show that this situation corresponds to the conditions

$$u_x(0, t) = 0, \quad u_x(l, t) = 0, \quad u(x, 0) = f(x).$$

Show that the method of separating variables yields the solution

$$u(x, t) = A_0 + \sum_{n=1}^{\infty} A_n \cos \frac{n\pi x}{l} e^{-(cn\pi/l)^2 t}$$

where, by (2) in Sec. 8.5,

$$A_0 = \frac{1}{l} \int_0^l f(x) dx, \quad A_n = \frac{2}{l} \int_0^l f(x) \cos \frac{n\pi x}{l} dx, \quad n = 1, 2, \dots.$$

13. In Prob. 12, $u \rightarrow A_0$ as $t \rightarrow \infty$. Does this agree with your physical intuition?

Find the temperature in the bar in Prob. 12, if $l = \pi$, $c = 1$, and

- | | | |
|--|--|----------------------------|
| 14. $f(x) = 1$ | 15. $f(x) = x^2$ | 16. $f(x) = \sin x$ |
| 17. $f(x) = \begin{cases} x & \text{when } 0 < x < \pi/2 \\ 0 & \text{when } \pi/2 < x < \pi \end{cases}$ | 18. $f(x) = \begin{cases} 1 & \text{when } 0 < x < \pi/2 \\ 0 & \text{when } \pi/2 < x < \pi \end{cases}$ | |

19. In Prob. 5, let the initial temperature be $u(x, 0) = f(x)$. Show that the temperature for any time $t > 0$ is $u(x, t) = u_I(x) + u_{II}(x, t)$ with u_I as before and

$$u_{II} = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{l} e^{-(cn\pi/l)^2 t},$$

$$B_n = \frac{2}{l} \int_0^l [f(x) - u_I(x)] \sin \frac{n\pi x}{l} dx = \frac{2}{l} \int_0^l f(x) \sin \frac{n\pi x}{l} dx + \frac{2}{n\pi} [(-1)^n U_2 - U_1].$$

20. Consider the bar in Probs. 6-11. Assume that the ends are kept at 100°C for a long time. Then at some instant, say, at $t = 0$, the temperature at $x = l$ is suddenly changed to 0°C and kept at this value, while the temperature at $x = 0$ is kept at 100°C . What are the temperatures in the middle of the bar at $t = 1, 2, 3, 10, 50$ sec?

9.6 Heat Flow in an Infinite Bar

We shall now consider solutions of the heat equation

$$(1) \quad \frac{\partial u}{\partial t} = c^2 \frac{\partial^2 u}{\partial x^2}$$

in the case of a bar which extends to infinity on both sides (and is laterally insulated, as before). In this case we do not have boundary conditions but only the initial condition

$$(2) \quad u(x, 0) = f(x) \quad (-\infty < x < \infty)$$

where $f(x)$ is the given initial temperature of the bar.

To solve our present problem we start as in the last section, that is, we substitute $u(x, t) = F(x)G(t)$ into (1). This yields the two ordinary differential equations

$$(3) \quad F'' + p^2 F = 0 \quad [\text{cf. (6), Sec. 9.5}]$$

and

$$(4) \quad G' + c^2 p^2 G = 0 \quad [\text{cf. (7), Sec. 9.5}].$$

The functions

$$F(x) = A \cos px + B \sin px \quad \text{and} \quad G(t) = e^{-c^2 p^2 t}$$

are solutions of (3) and (4), respectively; here A and B are arbitrary constants. Hence,

$$(5) \quad u(x, t; p) = FG = (A \cos px + B \sin px)e^{-c^2 p^2 t}$$

is a solution of (1). (As in the last section, we had to choose the constant of separation k negative, $k = -p^2$, because positive values of k lead to an increasing exponential function in (5), which has no physical meaning.)

Any series of functions (5), found in the usual manner by taking p as multiples of a fixed number, would lead to a function which is periodic in x

when $t = 0$. However, since $f(x)$ in (2) is not assumed periodic, it is natural to use Fourier integrals in the present case instead of Fourier series.

Since A and B in (5) are arbitrary, we may consider these quantities as functions of p and write $A = A(p)$ and $B = B(p)$. Since the heat equation is linear and homogeneous, the function

$$(6) \quad u(x, t) = \int_0^\infty u(x, t; p) dp = \int_0^\infty [A(p) \cos px + B(p) \sin px] e^{-c^2 p^2 t} dp$$

is then a solution of (1), provided this integral exists and can be differentiated twice with respect to x and once with respect to t .

From (6) and the initial condition (2) it follows that

$$(7) \quad u(x, 0) = \int_0^\infty [A(p) \cos px + B(p) \sin px] dp = f(x).$$

Using (4) and (5) in Sec. 8.10, we thus obtain

$$(8) \quad A(p) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(v) \cos pv dv, \quad B(p) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(v) \sin pv dv.$$

According to (16) in Prob. 20, Sec. 8.10, this Fourier integral may be written

$$u(x, 0) = \frac{1}{\pi} \int_0^\infty \left[\int_{-\infty}^{\infty} f(v) \cos(px - pv) dv \right] dp,$$

and (6), this section, thus becomes

$$u(x, t) = \frac{1}{\pi} \int_0^\infty \left[\int_{-\infty}^{\infty} f(v) \cos(px - pv) e^{-c^2 p^2 t} dv \right] dp.$$

Assuming that we may invert the order of integration, we obtain

$$(9) \quad u(x, t) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(v) \left[\int_0^\infty e^{-c^2 p^2 t} \cos(px - pv) dp \right] dv.$$

The inner integral can be evaluated by the use of the formula

$$(10) \quad \int_0^\infty e^{-s^2} \cos 2bs ds = \frac{\sqrt{\pi}}{2} e^{-b^2}$$

which will be derived in Sec. 15.4 (Prob. 29). Introducing a new variable of integration p by setting $s = cp\sqrt{t}$ and choosing

$$b = \frac{x - v}{2c\sqrt{t}}$$

we find that the formula (10) becomes

$$\int_0^\infty e^{-c^2 p^2 t} \cos(px - pv) dp = \frac{\sqrt{\pi}}{2c\sqrt{t}} e^{-(x-v)^2/4c^2 t}.$$

By inserting this result into (9) we obtain

$$(11) \quad u(x, t) = \frac{1}{2c\sqrt{\pi t}} \int_{-\infty}^{\infty} f(v) \exp \left\{ -\frac{(x-v)^2}{4c^2 t} \right\} dv.$$

Finally, by introducing the variable of integration $w = (v-x)/2c\sqrt{t}$, we see that this may be written

$$(12) \quad u(x, t) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} f(x + 2cw\sqrt{t}) e^{-w^2} dw.$$

If $f(x)$ is bounded for all values of x and integrable in every finite interval, it can be shown (cf. Ref. [E1]) that the function (11) or (12) satisfies (1) and (2). Hence this function is the required solution in the present case.

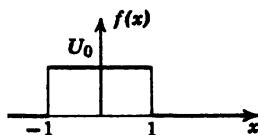


Fig. 256. Initial temperature in Example 1.

Example 1. Find the temperature in the infinite bar if the initial temperature is (Fig. 256)

$$f(x) = \begin{cases} U_0 = \text{const} & \text{when } |x| < 1, \\ 0 & \text{when } |x| > 1. \end{cases}$$

From (11) we have

$$u(x, t) = \frac{U_0}{2c\sqrt{\pi t}} \int_{-1}^1 \exp \left\{ -\frac{(x-v)^2}{4c^2 t} \right\} dv.$$

If we introduce the above variable of integration w , then the integration over v from -1 to 1 corresponds to the integration over w from $(-1-x)/2c\sqrt{t}$ to $(1-x)/2c\sqrt{t}$, and we thus obtain

$$(13) \quad u(x, t) = \frac{U_0}{\sqrt{\pi}} \int_{-(1+x)/2c\sqrt{t}}^{(1-x)/2c\sqrt{t}} e^{-w^2} dw \quad (t > 0).$$

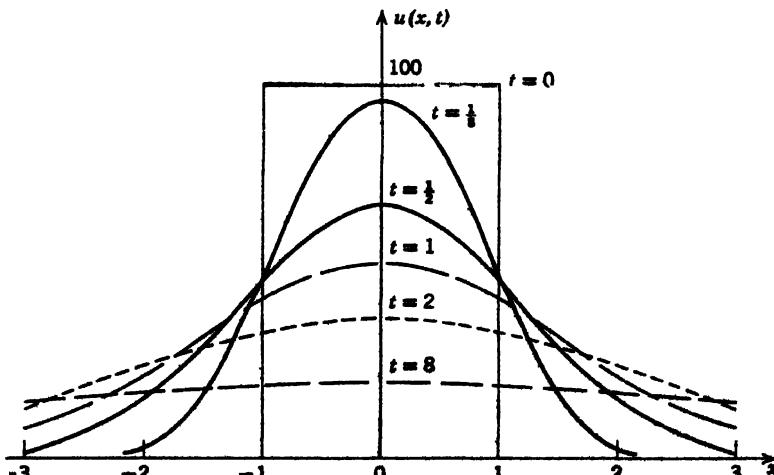


Fig. 257. Solution $u(x, t)$ of Example 1 for $U_0 = 100^\circ\text{C}$, $c^2 = 1 \text{ cm}^2/\text{sec}$, and several values of t .

We shall see later that this integral is not an elementary function, but can be easily expressed in terms of the so-called *error function*, whose values have been tabulated (cf. Sec. 17.2; cf. also Probs. 10–16, this section). Figure 257 shows $u(x, t)$ for $U_0 = 100^\circ\text{C}$, $c^2 = 1 \text{ cm}^2/\text{sec}$, and several values of t .

Problems

- Graph the temperatures in Ex. 1 (with $U_0 = 100^\circ\text{C}$ and $c^2 = 1 \text{ cm}^2/\text{sec}$) at the points $x = 0.5, 1$, and 1.5 as functions of t . Do the results agree with your physical intuition?
- If $f(x) = 1$ when $x > 0$ and $f(x) = 0$ when $x < 0$, show that (12) becomes

$$u(x, t) = \frac{1}{\sqrt{\pi}} \int_{-x/2c\sqrt{t}}^{\infty} e^{-w^2} dw \quad (t > 0).$$

- Show that for $x = 0$ the solution in Prob. 2 is independent of t . Can this result be expected for physical reasons?
- If the bar is semi-infinite, extending from 0 to ∞ , the end at $x = 0$ is held at temperature 0 and the initial temperature is $f(x)$, show that the temperature in the bar is

$$(14) \quad u(x, t) = \frac{1}{\sqrt{\pi}} \left[\int_{-x/\tau}^{\infty} f(x + \tau w) e^{-w^2} dw - \int_{x/\tau}^{\infty} f(-x + \tau w) e^{-w^2} dw \right], \quad \tau = 2c\sqrt{t}.$$

- Obtain (14) from (11) by assuming that $f(v)$ in (11) is odd.
- If in Prob. 4, $f(x) = 1$, show that

$$u(x, t) = \frac{2}{\sqrt{\pi}} \int_0^{x/\tau} e^{-w^2} dw \quad (t > 0).$$

- What form does (14) take if $f(x) = 1$ when $a < x < b$ (where $a > 0$) and $f(x) = 0$ otherwise?
- Show that the result of Prob. 6 can be obtained from (11) or (12) by using $f(x) = 1$ when $x > 0$, and $f(x) = -1$ when $x < 0$. What is the reason?
- Show that in Prob. 6 the times required for any two points to reach the same temperature are proportional to the squares of their distances from the boundary at $x = 0$.

A detailed treatment of the so-called error function

$$\operatorname{erf} x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-w^2} dw$$

will be given in Sec. 17.2. For the time being the student may get accustomed to this important function by solving the following problems.

- Show that $\operatorname{erf} x$ is odd.
- Show that

$$\int_a^b e^{-w^2} dw = \frac{\sqrt{\pi}}{2} (\operatorname{erf} b - \operatorname{erf} a), \quad \int_{-b}^b e^{-w^2} dw = \sqrt{\pi} \operatorname{erf} b.$$

- Using a table of the exponential function, graph the integrand of $\operatorname{erf} x$ (the so-called *bell-shaped curve*).

13. Obtain a small table of $\operatorname{erf} x$ for $x = 0, 0.2, 0.4, \dots, 1.0, 1.5, 2.0$ by counting squares under the curve in Prob. 12 or some other rough method of approximate integration, and compare the result with the actual values whose first two decimal places are 0.00, 0.22, 0.43, 0.60, 0.74, 0.84, 0.97, 1.00.
 14. Obtain the MacLaurin series of $\operatorname{erf} x$ by term-by-term integration of that of the integrand.
 15. Using the result of Prob. 14, compute $\operatorname{erf} x$ for the values in Prob. 13 (two decimal places).
 16. Show that (13) may be written
- $$u(x, t) = \frac{U_0}{2} \left[\operatorname{erf} \frac{1-x}{2c\sqrt{t}} + \operatorname{erf} \frac{1+x}{2c\sqrt{t}} \right] \quad (t > 0).$$
17. Using the formula in Prob. 16 and the table of the error function in Sec. 17.2, compute and graph $u(x, 1)$ in Ex. 1, assuming $U_0 = 100$ and $c = 1$. Compare with Fig. 257.
 18. Find the solution in Ex. 1 by starting from (12).
 19. Show that in Prob. 6, $u(x, t) = \operatorname{erf}(x/2c\sqrt{t})$ ($t > 0$).
 20. Graph u in Prob. 6 for $x \geq 0$ and (a) $t = 0$, (b) $t = 0.5$, (c) $t = 1$, and (d) $t = 2$. Choose $c = 1$.

9.7 Vibrating Membrane. Two-Dimensional Wave Equation

As another important problem from the field of vibrations, let us consider the motion of a stretched membrane, such as a drumhead. The reader will notice that our present considerations will be similar to those in the case of the vibrating string in Sec. 9.2.

We make the following assumptions:

1. *The mass of the membrane per unit area is constant ("homogeneous membrane"). The membrane is perfectly flexible and is so thin that it does not offer any resistance to bending.*
2. *The membrane is stretched and then fixed along its entire boundary in the xy -plane. The tension per unit length T caused by stretching the membrane is the same at all points and in all directions and does not change during the motion.*
3. *The deflection $u(x, y, t)$ of the membrane during the motion is small compared with the size of the membrane, and all angles of inclination are small.*

Although these assumptions cannot be realized in practice, small transverse vibrations of a thin physical membrane will satisfy these assumptions relatively accurately.

To derive the differential equation which governs the motion of the membrane, we consider the forces acting on a small portion of the membrane as shown in Fig. 258. Since the deflections of the membrane and the angles

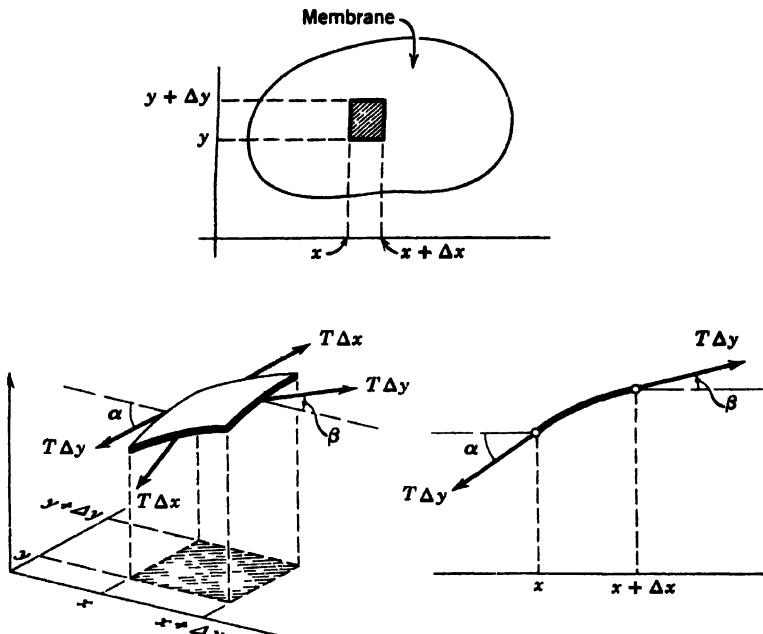


Fig. 258. Vibrating membrane.

of inclination are small, the sides of the portion are approximately equal to Δx and Δy . The tension T is the force per unit length. Hence the forces acting on the edges of the portion are approximately $T \Delta x$ and $T \Delta y$. Since the membrane is perfectly flexible, these forces are tangent to the membrane.

We first consider the horizontal components of the forces. These components are obtained by multiplying the forces by the cosines of the angles of inclination. Since these angles are small, their cosines are close to 1. Hence the horizontal components of the forces at opposite edges are approximately equal. Therefore, the motion of the particles of the membrane in horizontal direction will be negligibly small. From this we conclude that we may regard the motion of the membrane as transversal, that is, each particle moves vertically.

The vertical components of the forces along the edges parallel to the y -plane are⁴ (Fig. 258)

$$T \Delta y \sin \beta \quad \text{and} \quad -T \Delta y \sin \alpha;$$

here the minus sign appears because the force on the left edge is directed downward. Since the angles are small, we may replace their sines by their tangents. Hence the resultant of those two vertical components is

$$(1) \quad T \Delta y(\sin \beta - \sin \alpha) \approx T \Delta y(\tan \beta - \tan \alpha) \\ = T \Delta y[u_x(x + \Delta x, y_1) - u_x(x, y_2)]$$

⁴ Note that the angle of inclination varies along the edges, and α and β represent values of that angle at a suitable point of the edges under consideration.

where subscripts x denote partial derivatives and y_1 and y_2 are values between y and $y + \Delta y$. Similarly, the resultant of the vertical components of the forces acting on the other two edges of the portion is

$$(2) \quad T \Delta x [u_v(x_1, y + \Delta y) - u_v(x_2, y)]$$

where x_1 and x_2 are values between x and $x + \Delta x$.

By Newton's second law (cf. Sec. 2.6), the sum of the forces given by (1) and (2) is equal to the mass $\rho \Delta A$ of the portion times the acceleration $\partial^2 u / \partial t^2$; here ρ is the mass of the undeflected membrane per unit area and $\Delta A = \Delta x \Delta y$ is the area of the portion when it is undeflected. Thus

$$\begin{aligned} \rho \Delta x \Delta y \frac{\partial^2 u}{\partial t^2} &= T \Delta y [u_x(x + \Delta x, y_1) - u_x(x, y_2)] \\ &\quad + T \Delta x [u_v(x_1, y + \Delta y) - u_v(x_2, y)] \end{aligned}$$

where the derivative on the left is evaluated at some suitable point (\tilde{x}, \tilde{y}) corresponding to the portion. Division by $\rho \Delta x \Delta y$ yields

$$\frac{\partial^2 u}{\partial t^2} = \frac{T}{\rho} \left[\frac{u_x(x + \Delta x, y_1) - u_x(x, y_2)}{\Delta x} + \frac{u_v(x_1, y + \Delta y) - u_v(x_2, y)}{\Delta y} \right]$$

If we let Δx and Δy approach zero, then we obtain

$$(3) \quad \frac{\partial^2 u}{\partial t^2} = c^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \quad c^2 = \frac{T}{\rho}.$$

This equation is called the **two-dimensional wave equation**. We see that the expression in parentheses is the Laplacian $\nabla^2 u$ of u (cf. Sec. 5.16) and we may write (3) in the form

$$(3') \quad \frac{\partial^2 u}{\partial t^2} = c^2 \nabla^2 u.$$

9.8 Rectangular Membrane

To solve the problem of a vibrating membrane, we have to determine a solution $u(x, y, t)$ of the two-dimensional wave equation

$$(1) \quad \frac{\partial^2 u}{\partial t^2} = c^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right)$$

that satisfies the boundary condition

$$(2) \quad u = 0 \quad \text{on the boundary of the membrane for all } t \geq 0$$

and the initial conditions

$$(3) \quad u(x, y, 0) = f(x, y) \quad [\text{given initial displacement } f(x, y)]$$

and

$$(4) \quad \left. \frac{\partial u}{\partial t} \right|_{t=0} = g(x, y) \quad [\text{given initial velocity } g(x, y)].$$

These conditions are quite similar to those in the case of the vibrating string.

As a first important case, let us consider the rectangular membrane R shown in Fig. 259.

First step. By applying the method of separation of variables we first determine solutions of (1) that satisfy the condition

(2). For this purpose we start from

$$(5) \quad u(x, y, t) = F(x, y) G(t).$$

By substituting this into the wave equation (1) we have

$$\ddot{FG} = c^2(F_{xx}G + F_{yy}G)$$

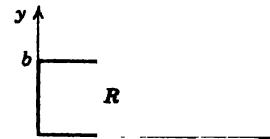


Fig. 259. Rectangular membrane.

where subscripts denote partial derivatives and dots denote derivatives with respect to t . Dividing both sides by c^2FG , we find

$$\frac{\ddot{G}}{c^2G} = \frac{1}{F}(F_{xx} + F_{yy}).$$

Since the functions on the left depend only on t while the functions on the right do not depend on t , the expressions on both sides must be equal to a constant. A little investigation shows that only negative values of that constant will lead to solutions which satisfy (2) without being identically zero; this is similar to the procedure in Sec. 9.3. Denoting this negative constant by $-\nu^2$, we thus have

$$\frac{\ddot{G}}{c^2G} = \frac{1}{F}(F_{xx} + F_{yy}) = -\nu^2.$$

This yields the two differential equations

$$(6) \quad \ddot{G} + \lambda^2G = 0 \quad \text{where} \quad \lambda = c\nu,$$

and

$$(7) \quad F_{xx} + F_{yy} + \nu^2F = 0.$$

Now we consider (7) and apply the method of separating variables once more, that is, we determine solutions of (7) of the form

$$(8) \quad F(x, y) = H(x)Q(y)$$

which are zero on the boundary of the membrane. Substitution of (8) into (7) yields

$$\frac{d^2H}{dx^2} Q = - \left(H \frac{d^2Q}{dy^2} + \nu^2HQ \right).$$

By dividing both sides by HQ we find

$$\frac{1}{H} \frac{d^2H}{dx^2} = - \frac{1}{Q} \left(\frac{d^2Q}{dy^2} + \nu^2Q \right).$$

The functions on the left depend only on x whereas the functions on the right depend only on y . Hence the expressions on both sides must be equal to a constant. This constant must be negative, say, $-k^2$, because only negative values will lead to solutions that satisfy (2) without being identically zero. Thus

$$\frac{1}{H} \frac{d^2H}{dx^2} = -\frac{1}{Q} \left(\frac{d^2Q}{dy^2} + v^2 Q \right) = -k^2.$$

This yields the ordinary differential equations

$$(9) \quad \frac{d^2H}{dx^2} + k^2 H = 0$$

and

$$(10) \quad \frac{d^2Q}{dy^2} + p^2 Q = 0 \quad \text{where } p^2 = v^2 - k^2.$$

Second step. The general solutions of (9) and (10) are

$$H(x) = A \cos kx + B \sin kx \quad \text{and} \quad Q(y) = C \cos py + D \sin py$$

where A , B , C , and D are constants. From (5) and (2) it follows that $F = HQ$ must be zero on the boundary, which corresponds to $x = 0$, $x = a$, $y = 0$, and $y = b$; cf. Fig. 259. This yields the conditions

$$H(0) = 0, \quad H(a) = 0, \quad Q(0) = 0, \quad Q(b) = 0.$$

Therefore, $H(0) = A = 0$, and then

$$H(a) = B \sin ka = 0.$$

We must take $B \neq 0$ since otherwise $H \equiv 0$ and $F \equiv 0$. Hence $\sin ka = 0$ or $ka = m\pi$, that is,

$$k = \frac{m\pi}{a} \quad (m \text{ integral}).$$

In precisely the same fashion we conclude that $C = 0$ and p must be restricted to the values $p = n\pi/b$ where n is integral. In this way we obtain the solutions

$$H_m(x) = \sin \frac{m\pi x}{a} \quad \text{and} \quad Q_n(y) = \sin \frac{n\pi y}{b} \quad \begin{matrix} m = 1, 2, \dots, \\ n = 1, 2, \dots \end{matrix}$$

(As in the case of the vibrating string, it is not necessary to consider $m, n = -1, -2, \dots$ because the corresponding solutions are essentially the same as for positive m and n , except for a factor -1 .) It follows that the functions

$$(11) \quad F_{mn}(x, y) = H_m(x) Q_n(y) = \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} \quad \begin{matrix} m = 1, 2, \dots, \\ n = 1, 2, \dots \end{matrix}$$

are solutions of the equation (7) which are zero on the boundary of the rectangular membrane.

Since $p^2 = \nu^2 - k^2$ in (10) and $\lambda = cv$ in (6), we have

$$\lambda = c\sqrt{k^2 + p^2}.$$

Hence to $k = m\pi/a$ and $p = n\pi/b$ there corresponds the value

$$(12) \quad \lambda = \lambda_{mn} = c\pi\sqrt{\frac{m^2}{a^2} + \frac{n^2}{b^2}} \quad m = 1, 2, \dots, \\ n = 1, 2, \dots,$$

in (6), and the corresponding general solution of (6) is

$$G_{mn}(t) = B_{mn} \cos \lambda_{mn} t + B_{mn}^* \sin \lambda_{mn} t.$$

It follows that the functions $u_{mn}(x, y, t) = F_{mn}(x, y)G_{mn}(t)$, at length

$$(13) \quad u_{mn}(x, y, t) = (B_{mn} \cos \lambda_{mn} t + B_{mn}^* \sin \lambda_{mn} t) \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b}$$

with λ_{mn} according to (12) are solutions of the wave equation (1), which are zero on the boundary of the rectangular membrane in Fig. 259. These functions are called the **eigenfunctions** or *characteristic functions*, and the numbers λ_{mn} are called the **eigenvalues** or *characteristic values* of the vibrating membrane. The frequency of u_{mn} is $\lambda_{mn}/2\pi$.

It is interesting to note that, depending on a and b , several functions F_{mn} may correspond to the same eigenvalue. Physically this means that there may exist vibrations having the same frequency but entirely different nodal lines (curves of points on the membrane which do not move). This may be illustrated by the following example.

Example 1. Consider the square membrane for which $a = b = 1$. From (12) we see that the eigenvalues are

$$(14) \quad \lambda_{mn} = c\pi\sqrt{m^2 + n^2}.$$

Hence,

$$\lambda_{mn} = \lambda_{nm}$$

but for $m \neq n$ the corresponding functions

$$F_{mn} = \sin m\pi x \sin n\pi y \quad \text{and} \quad F_{nm} = \sin n\pi x \sin m\pi y$$

are certainly different. For example, to $\lambda_{12} = \lambda_{21} = c\pi\sqrt{5}$ there correspond the two functions

$$F_{12} = \sin \pi x \sin 2\pi y \quad \text{and} \quad F_{21} = \sin 2\pi x \sin \pi y.$$

Hence the corresponding solutions

$$u_{12} = (B_{12} \cos c\pi\sqrt{5}t + B_{12}^* \sin c\pi\sqrt{5}t)F_{12}$$

and

$$u_{21} = (B_{21} \cos c\pi\sqrt{5}t + B_{21}^* \sin c\pi\sqrt{5}t)F_{21}$$

have the nodal lines $y = 1/2$ and $x = 1/2$, respectively (cf. Fig. 260 on the next page). Taking $B_{12} = 1$ and $B_{12}^* = B_{21}^* = 0$, we obtain

$$(15) \quad u_{12} + u_{21} = \cos c\pi\sqrt{5}t (F_{12} + B_{21}F_{21}).$$

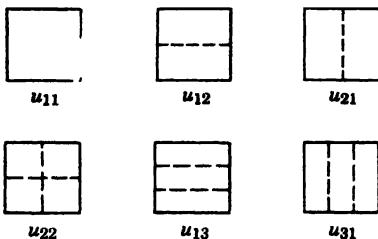


Fig. 260. Nodal lines of the solutions u_{11} , u_{12} , u_{21} , u_{22} , u_{13} , and u_{31} in the case of the square membrane.

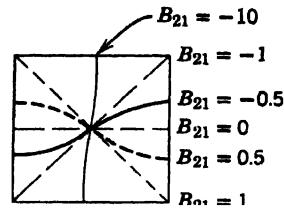


Fig. 261. Nodal lines of the solution (15) for some values of B_{21} .

which represents another vibration corresponding to the eigenvalue $c\pi\sqrt{5}$. The nodal line of this function is the solution of the equation

$$F_{12} + B_{21}F_{21} = \sin \pi x \sin 2\pi y + B_{21} \sin 2\pi x \sin \pi y = 0$$

or, since $\sin 2\alpha = 2 \sin \alpha \cos \alpha$,

$$(16) \quad \sin \pi x \sin \pi y (\cos \pi y + B_{21} \cos \pi x) = 0.$$

This solution depends on the value of B_{21} (Fig. 261).

From (14) we see that even more than two functions may correspond to the same numerical value of λ_{mn} . For example, the four functions F_{18} , F_{81} , F_{47} , and F_{74} correspond to $\lambda_1 = \lambda_{47} = \lambda_{74} = c\pi\sqrt{65}$, because

$$1^2 + 8^2 = 4^2 + 7^2 = 65.$$

This happens because 65 can be expressed as the sum of two squares of natural numbers in more ways than one way. According to a theorem by Gauss, this is the case for every sum of two squares among whose prime factors there are at least two different ones of the form $4n + 1$ where n is a positive integer. In our case,

$$65 = 5 \cdot 13 = (4 + 1)(12 + 1).$$

Third step. To obtain the solution which also satisfies the initial conditions (3) and (4), we proceed in a similar fashion as in Sec. 9.3. We consider the double series⁵

$$(17) \quad \begin{aligned} u(x, y, t) &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} u_{mn}(x, y, t) \\ &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} (B_{mn} \cos \lambda_{mn} t + B_{mn}^* \sin \lambda_{mn} t) \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} \end{aligned}$$

From this and (3) we obtain

$$(18) \quad u(x, y, 0) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} B_{mn} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} = f(x, y).$$

This series is called a **double Fourier series**. Suppose that $f(x, y)$ can be developed in such a series.⁶ Then the Fourier coefficients B_{mn} of $f(x, y)$

⁵ We shall not consider the problems of convergence and uniqueness.

⁶ Sufficient conditions: f , $\partial f / \partial x$, $\partial f / \partial y$, $\partial^2 f / \partial x \partial y$ continuous in the rectangle R under consideration.

in (18) can be determined as follows. Setting

$$(19) \quad K_m(y) = \sum_{n=1}^{\infty} B_{mn} \sin \frac{n\pi y}{b}$$

we may write (18) in the form

$$f(x, y) = \sum_{m=1}^{\infty} K_m(y) \sin \frac{m\pi x}{a}.$$

For fixed y this is the Fourier sine series of $f(x, y)$, considered as a function of x , and from (4) in Sec. 8.5 it follows that the coefficients of this expansion are

$$(20) \quad K_m(y) = \frac{2}{a} \int_0^a f(x, y) \sin \frac{m\pi x}{a} dx.$$

Furthermore, (19) is the Fourier sine series of $K_m(y)$, and from (4) in Sec. 8.5 it follows that the coefficients are

$$B_{mn} = \frac{2}{b} \int_0^b K_m(y) \sin \frac{n\pi y}{b} dy.$$

From this and (20) we obtain the *generalized Euler formula*

$$(21) \quad B_{mn} = \frac{4}{ab} \int_0^b \int_0^a f(x, y) \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} dx dy \quad m = 1, 2, \dots \quad n = 1, 2, \dots$$

for the Fourier coefficients of $f(x, y)$ in the double Fourier series (18).

The B_{mn} in (17) are now determined in terms of $f(x, y)$. To determine the B_{mn}^* , we differentiate (17) termwise with respect to t ; using (4), we obtain

$$\left. \frac{\partial u}{\partial t} \right|_{t=0} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} B_{mn}^* \lambda_{mn} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} = g(x, y).$$

Suppose that $g(x, y)$ can be developed in this double Fourier series. Then, proceeding as before, we find

$$(22) \quad B_{mn}^* = \frac{4}{ab \lambda_{mn}} \int_0^b \int_0^a g(x, y) \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b} dx dy \quad m = 1, 2, \dots \quad n = 1, 2, \dots$$

The result is that, for (17) to satisfy the initial conditions, the coefficients B_{mn} and B_{mn}^* must be chosen according to (21) and (22).

Problems

- How does the frequency of a solution (13) change, if the tension of the membrane is increased?
- Determine and graph the nodal lines of the solutions (13) with $m = 1, 2, 3, 4$ and $n = 1, 2, 3, 4$ in the case $a = b = 1$.
- Same task as in Prob. 2, when $a = 2$ and $b = 1$.
- Find further eigenvalues of the square membrane with side 1 such that four different eigenfunctions correspond to each such eigenvalue.

5. Find eigenvalues of the rectangular membrane of sides $a = 2, b = 1$ such that two or more different eigenfunctions correspond to each such eigenvalue.
6. Show that, among all rectangular membranes of the same area $A = ab$ and the same c , the square membrane is that for which u_{11} [cf. (13)] has the lowest frequency.
7. Find a similar result as in Prob. 6 for the frequency of any solution (13) (m, n fixed).

Represent the following functions $f(x, y)$ ($0 < x < a, 0 < y < b$) by a double Fourier series of the form (18).

8. $f = 1$

9. $f = x + y$

10. $f = xy$

11. $f = xy(a - x)(b - y)$

12. $f = xy(a^2 - x^2)(b^2 - y^2)$

13. Show that

$$u_{tt} = c^2 \nabla^2 u + P/\rho$$

governs the forced vibrations of the membrane, where $P(x, y, t)$ is the external force per unit area acting normal to the xy -plane.

14. Find the eigenfunctions of the rectangular membrane in Fig. 262 which is fixed on the boundary.

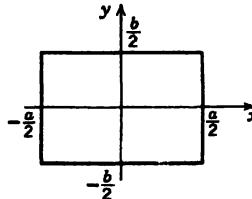


Fig. 262. Rectangular membrane.

Find the deflection $u(x, y, t)$ of the square membrane with $a = b = 1$ and $c = 1$, if the initial velocity is zero and the initial deflection is $f(x, y)$, where

15. $f = 0.01xy(1 - x)(1 - y)$

16. $f = kx(1 - x^2)y(1 - y^2)$

17. $f = k \sin \pi x \sin 2\pi y$

18. $f = k \sin^2 \pi x \sin^2 \pi y$

19. The edges of a thin square plate (Fig. 263) are kept at temperature zero and the faces are perfectly insulated. The initial temperature is assumed to be $u(x, y, 0) = f(x, y)$. By applying the method of separating variables to the two-dimensional heat equation $u_t = c^2 \nabla^2 u$ show that the temperature in the plate is

$$u(x, y, t) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} B_{mn} \sin mx \sin ny e^{-c^2(m^2+n^2)t}$$

where

$$B_{mn} = \frac{4}{\pi^2} \int_0^\pi \int_0^\pi f(x, y) \sin mx \sin ny dx dy.$$

20. Find the temperature in the plate of Prob. 19, if $f(x, y) = x(\pi - x)y(\pi - y)$.

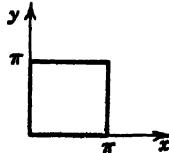


Fig. 263. Square plate.

9.9 Laplacian in Polar Coordinates

In connection with boundary value problems for partial differential equations, it is a general principle to use coordinates with respect to which the boundary of the region under consideration has a simple representation. In the next section we shall consider circular membranes. Then polar coordinates r and θ , defined by

$$x = r \cos \theta, \quad y = r \sin \theta,$$

will be appropriate, because the boundary of the membrane can then be represented by the simple equation $r = \text{const.}$

When using r and θ we have to transform the Laplacian

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$$

in the wave equation into these new coordinates.

Transformations of differential expressions from one coordinate system into another are frequently required in applications. Therefore, the student should follow our present consideration with great attention.

As in Sec. 9.4, we shall use the chain rule. For the sake of simplicity we shall denote partial derivatives by subscripts and $u(x, y, t)$, as a function of r, θ, t , by the same letter u .

By applying the chain rule (4), Sec. 5.15, we obtain

$$u_x = u_r r_x + u_\theta \theta_x.$$

By differentiating this again with respect to x we first have

$$(1) \quad \begin{aligned} u_{xx} &= (u_r r_x)_x + (u_\theta \theta_x)_x \\ &= (u_r)_x r_x + u_r r_{xx} + (u_\theta)_x \theta_x + u_\theta \theta_{xx}. \end{aligned}$$

Now, by applying the chain rule again, we find

$$(u_r)_x = u_{rr} r_x + u_{r\theta} \theta_x \quad \text{and} \quad (u_\theta)_x = u_{\theta r} r_x + u_{\theta\theta} \theta_x.$$

To determine the partial derivatives r_x and θ_x , we have to differentiate

$$r = \sqrt{x^2 + y^2} \quad \text{and} \quad \theta = \arctan \frac{y}{x},$$

finding

$$r_x = \frac{x}{\sqrt{x^2 + y^2}} = \frac{x}{r}, \quad \theta_x = \frac{1}{1 + (y/x)^2} \left(-\frac{y}{x^2} \right) = -\frac{y}{r^2}.$$

Differentiating these two formulas again, we obtain

$$r_{xx} = \frac{r - xr_x}{r^2} = \frac{1}{r} - \frac{x^2}{r^3} = \frac{y^2}{r^3}, \quad \theta_{xx} = -y \left(-\frac{2}{r^3} \right) r_x = \frac{2xy}{r^4}.$$

Substituting all these expressions into (1), using $u_{r\theta} = u_{\theta r}$ (cf. Sec. 0.2), and simplifying, we readily find

$$(2) \quad u_{xx} = \frac{x^2}{r^2} u_{rr} - 2 \frac{xy}{r^3} u_{r\theta} + \frac{y^2}{r^4} u_{\theta\theta} + \frac{y^2}{r^3} u_r + 2 \frac{xy}{r^4} u_\theta.$$

In a similar fashion we obtain

$$(3) \quad u_{yy} = \frac{y^2}{r^2} u_{rr} + 2 \frac{xy}{r^3} u_{r\theta} + \frac{x^2}{r^4} u_{\theta\theta} + \frac{x^2}{r^3} u_r - 2 \frac{xy}{r^4} u_\theta.$$

By adding these two expressions we see that the Laplacian in polar coordinates is

$$(4) \quad \nabla^2 u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2}.$$

Problems

1. Show that (4) may be written

$$\nabla^2 u = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2}.$$

2. If u is independent of θ , then (4) reduces to $\nabla^2 u = u_{rr} + u_r/r$. Derive this result directly from the Laplacian in Cartesian coordinates by assuming that u is independent of θ .

3. Transform (4) back into Cartesian coordinates.

4. If x, y are Cartesian coordinates, show that $x^* = x \cos \alpha - y \sin \alpha$, $y^* = x \sin \alpha + y \cos \alpha$ are Cartesian coordinates, and verify by calculation that $\nabla^2 u = u_{x^*x^*} + u_{y^*y^*}$.

5. Express $\nabla^2 u$ in terms of the coordinates $x^* = ax + b$, $y^* = cy + d$, where x, y are Cartesian coordinates.

6. Find the Laplacian in cylindrical coordinates r, θ, z defined by $x = r \cos \theta$, $y = r \sin \theta$, $z = z$.

7. Let r, θ, ϕ be spherical coordinates, defined by

$$x = r \cos \theta \sin \phi, \quad y = r \sin \theta \sin \phi, \quad z = r \cos \phi.$$

If $u(x, y, z)$ is a function of $r = \sqrt{x^2 + y^2 + z^2}$ only, show that

$$\nabla^2 u = u_{rr} + \frac{2}{r} u_r.$$

8. Show that the Laplacian of u in spherical coordinates is

$$\nabla^2 u = u_{rr} + \frac{2}{r} u_r + \frac{1}{r^2} u_{\phi\phi} + \frac{\cot \phi}{r^2} u_\phi + \frac{1}{r^2 \sin^2 \phi} u_{\theta\theta}.$$

9. If the surface of the homogeneous solid sphere $x^2 + y^2 + z^2 \leq R^2$ is kept at temperature zero and the initial temperature in the sphere is $f(r)$ where $r = \sqrt{x^2 + y^2 + z^2}$, show that the temperature $u(r, t)$ in the sphere is the solution of $u_t = c^2 \left(u_{rr} + \frac{2}{r} u_r \right)$, satisfying the conditions $u(R, t) = 0$, $u(r, 0) = f(r)$.

10. Show that by setting $v = ru$ the formulas in Prob. 9 take the form $v_t = c^2 v_{rr}$, $v(R, t) = 0$, $v(r, 0) = rf(r)$. Include the condition $v(0, t) = 0$ (which holds because u must be bounded at $r = 0$), and solve the resulting problem by separating variables.

9.10 Circular Membrane. Bessel's Equation

We shall now consider vibrations of the circular membrane of radius R (Fig. 264). Using polar coordinates defined by $x = r \cos \theta$, $y = r \sin \theta$, the wave equation (3') in Sec. 9.7 takes the form

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} \right),$$

cf. (4) in Sec. 9.9. In the present section we shall consider solutions $u(r, t)$ of this equation which are radially symmetric, that is, do not depend on θ . Then the wave equation reduces to the simpler form

$$(1) \quad \frac{\partial^2 u}{\partial t^2} = c^2 \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} \right).$$

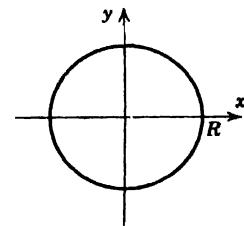


Fig. 264. Circular membrane.

Since the membrane is fixed along the boundary $r = R$, we have the boundary condition

$$(2) \quad u(R, t) = 0 \quad \text{for all } t \geq 0.$$

Solutions not depending on θ will occur if the initial conditions do not depend on θ , that is, if they are of the form

$$(3) \quad u(r, 0) = f(r) \quad [\text{initial deflection } f(r)]$$

and

$$(4) \quad \left. \frac{\partial u}{\partial t} \right|_{t=0} = g(r) \quad [\text{initial velocity } g(r)].$$

First step. Using the method of separating variables, we first determine solutions of (1) that satisfy the boundary condition (2). We start from

$$(5) \quad u(r, t) = W(r)G(t).$$

By differentiating and inserting (5) into (1) and dividing the resulting equation by $c^2 W G$ we obtain

$$\frac{\ddot{G}}{c^2 G} = \frac{1}{W} (W'' + \frac{1}{r} W')$$

where dots denote derivatives with respect to t and primes denote derivatives with respect to r . The expressions on both sides must be equal to a constant,

and this constant must be negative, say, $-k^2$, in order to obtain solutions that satisfy the boundary condition without being identically zero. Thus,

$$\frac{\ddot{G}}{c^2 G} = \frac{1}{W} \left(W'' + \frac{1}{r} W' \right) = -k^2.$$

This yields the ordinary differential equations

$$(6) \quad \ddot{G} + \lambda^2 G = 0 \quad \text{where } \lambda = ck$$

and

$$(7) \quad W'' + \frac{1}{r} W' + k^2 W = 0$$

Second step. We first consider (7). Introducing the new independent variable $s = kr$ we have $1/r = k/s$,

$$W' = \frac{dW}{dr} = \frac{dW}{ds} \frac{ds}{dr} = \frac{dW}{ds} k \quad \text{and} \quad W'' = \frac{d^2 W}{ds^2} k^2.$$

By substituting this into (7) and omitting the common factor k^2 we obtain

$$\frac{d^2 W}{ds^2} + \frac{1}{s} \frac{dW}{ds} + W = 0.$$

This is **Bessel's equation** (1), Sec. 3.5, with $\nu = 0$. The general solution is (cf. Sec. 3.7)

$$W = C_1 J_0(s) + C_2 Y_0(s)$$

where J_0 and Y_0 are the Bessel functions of the first and second kind of order zero. Since the deflection of the membrane is always finite while Y_0 becomes infinite as s approaches zero, we cannot use Y_0 and must choose $C_2 = 0$. Clearly $C_1 \neq 0$ since otherwise $W \equiv 0$. We may set $C_1 = 1$, and then

$$(8) \quad W(r) = J_0(s) = J_0(kr).$$

On the boundary $r = R$ we must have $u(R, t) = W(R)G(t) = 0$. Since $G \equiv 0$ would imply $u \equiv 0$, we require that

$$W(R) = J_0(kR) = 0.$$

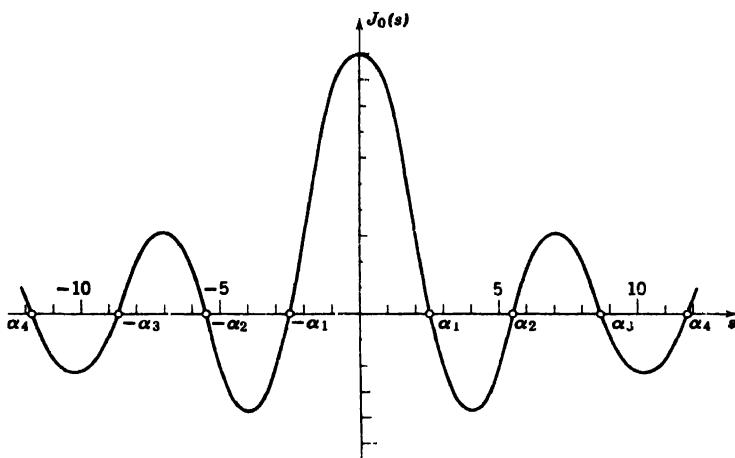
Denoting the positive zeros of $J_0(s)$ by $s = \alpha_1, \alpha_2, \dots$ (Fig. 265), we thus obtain

$$(9) \quad kR = \alpha_m \quad \text{or} \quad k = k_m = \frac{\alpha_m}{R}, \quad m = 1, 2, \dots$$

Hence the functions

$$(10) \quad W_m(r) = J_0(k_m r) = J_0\left(\frac{\alpha_m}{R} r\right) \quad m = 1, 2, \dots$$

are solutions of (7) which vanish at $r = R$.

Fig. 265. Bessel function $J_0(s)$.

The corresponding general solutions of (6) with $\lambda = \lambda_m = ck_m$ are

$$G_m(t) = a_m \cos \lambda_m t + b_m \sin \lambda_m t.$$

Hence the functions

$$(11) \quad u_m(r, t) = W_m(r)G_m(t) = (a_m \cos \lambda_m t + b_m \sin \lambda_m t)J_0(k_m r)$$

where $m = 1, 2, \dots$, are solutions of the wave equation (1), satisfying the boundary condition (2). These are the *eigenfunctions* of our problem, and the corresponding *eigenvalues* are λ_m .

The vibration of the membrane corresponding to u_m is called the *mth normal mode*; it has the frequency $\lambda_m/2\pi$ cycles per unit time. Since the zeros of J_0 are not regularly spaced on the axis (in contrast to the zeros of the sine functions appearing in the case of the vibrating string), the sound of a drum is entirely different from that of a violin. The forms of the normal modes can easily be obtained from Fig. 265 and are shown in Fig. 266.

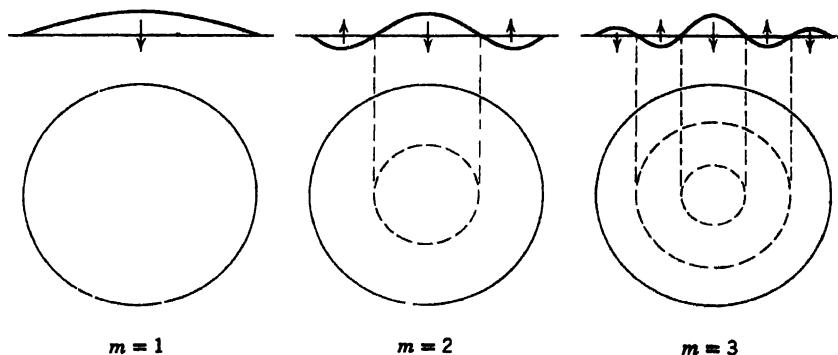


Fig. 266. Normal modes of the circular membrane in the case of vibrations independent of the angle.

Table 22. Zeros $s = \alpha_m$ of $J_0(s)$.

m	α_m	m	α_m	m	α_m
1	2.404 825 6	11	33.775 820 2	21	65.189 964 8
2	5.520 078 1	12	36.917 098 4	22	68.331 469 3
3	8.653 727 9	13	40.058 425 8	23	71.472 981 6
4	11.791 534 4	14	43.199 791 7	24	74.614 500 6
5	14.930 917 7	15	46.341 188 4	25	77.756 025 6
6	18.071 064 0	16	49.482 609 9	26	80.897 555 9
7	21.211 636 6	17	52.624 051 8	27	84.039 090 8
8	24.352 471 5	18	55.765 510 8	28	87.180 629 8
9	27.493 479 1	19	58.906 983 9	29	90.322 172 6
10	30.634 606 5	20	62.048 469 2	30	93.463 718 8

For $m = 1$, all the points of the membrane move upward (or downward) at the same time. For $m = 2$, the situation is as follows. The function

$$W_2(r) = J_0\left(\frac{\alpha_2}{R} r\right)$$

is zero for $\alpha_2 r/R = \alpha_1$ or $r = \alpha_1 R/\alpha_2$. The circle $r = \alpha_1 R/\alpha_2$ is, therefore, a nodal line, and when at some instant the central part of the membrane moves upward, the outer part ($r > \alpha_1 R/\alpha_2$) moves downward, and conversely. The solution $u_m(r, t)$ has $m - 1$ nodal lines which are concentric circles (Fig. 266).

Third step. To obtain a solution that also satisfies the initial conditions (3) and (4), we may proceed as in the case of the vibrating string, that is, we consider the series⁷

$$(12) \quad u(r, t) = \sum_{m=1}^{\infty} W_m(r) G_m(t) = \sum_{m=1}^{\infty} (a_m \cos \lambda_m t + b_m \sin \lambda_m t) J_0\left(\frac{\alpha_m}{R} r\right).$$

Setting $t = 0$ and using (3), we obtain

$$(13) \quad u(r, 0) = \sum_{m=1}^{\infty} a_m J_0\left(\frac{\alpha_m}{R} r\right) = f(r).$$

Hence, for (12) to satisfy (3), the a_m must be the coefficients of the Fourier-Bessel series which represents $f(r)$ in terms of $J_0(\alpha_m r/R)$, that is [cf. (9) in Sec. 8.13],

$$= \frac{2}{R^2 J_1^2(\alpha_m)} \int_0^R r f(r) J_0\left(\frac{\alpha_m}{R} r\right) dr, \quad m = 1, 2,$$

Differentiability of $f(r)$ in the interval $0 \leq r \leq R$ is sufficient for the existence of the development (13), cf. Ref. [B9]. The coefficients b_m in (12) can be determined from (4) in a similar fashion. To obtain numerical values of a_m and b_m , we may apply one of the usual methods of approximate integration, using tables of J_0 and J_1 .

⁷ We shall not consider the problems of convergence and uniqueness.

Problems

1. Determine numerical values of the radii of the nodal lines of u_2 and u_3 [cf. (11)] when $R = 1$.
2. Same question as in Prob. 1, for u_4 , u_5 , and u_6 .
3. Sketch a figure similar to Fig. 266, for u_4 , u_5 , u_6 .
4. If the tension of the membrane is increased how does the frequency of each normal mode (11) change?
5. Show that for (12) to satisfy (4),

$$b_m = \frac{2}{c\alpha_m R J_1^2(\alpha_m)} \int_0^R r g(r) J_0(\alpha_m r/R) dr, \quad m = 1, 2, \dots.$$

6. Is it possible that, for fixed c and R , two or more functions u_m [cf. (11)] that have different nodal lines correspond to the same eigenvalue?
7. (Vibrations depending on r and θ) Show that substitution of $u = F(r, \theta)G(t)$ into the wave equation

$$(14) \qquad u_{tt} = c^2 \left(u_{rr} + \frac{1}{r} u_r + \frac{1}{r^2} u_{\theta\theta} \right)$$

leads to

$$(15) \qquad \ddot{G} + \lambda^2 G = 0, \quad \text{where } \lambda = ck,$$

$$(16) \qquad F_{rr} + \frac{1}{r} F_r + \frac{1}{r^2} F_{\theta\theta} + k^2 F = 0.$$

8. Show that substitution of $F = W(r)Q(\theta)$ in (16) yields

$$(17) \qquad Q'' + n^2 Q = 0,$$

$$(18) \qquad r^2 W'' + r W' + (k^2 r^2 - n^2) W = 0.$$

9. Show that $Q(0)$ must be periodic with period 2π , and, therefore, $n = 0, 1, \dots$ in (17) and (18). Show that this yields the solutions $Q_n = \cos n\theta$, $Q_n^* = \sin n\theta$, $W_n = J_n(kr)$, $n = 0, 1, \dots$.

10. Show that the boundary condition

$$(19) \qquad u(R, \theta, t) = 0$$

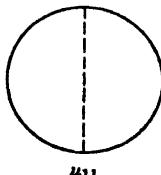
leads to the values $k = k_{mn} = \alpha_{mn}/R$ where $s = \alpha_{mn}$ is the m th positive zero of $J_n(s)$.

11. Show that solutions of (14) which satisfy (19) are

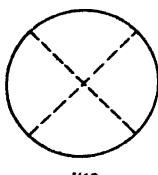
$$(20) \qquad \begin{aligned} u_{mn} &= (A_{mn} \cos ck_{mn}t + B_{mn} \sin ck_{mn}t) J_n(k_{mn}r) \cos n\theta, \\ u_{mn}^* &= (A_{mn}^* \cos ck_{mn}t + B_{mn}^* \sin ck_{mn}t) J_n(k_{mn}r) \sin n\theta. \end{aligned}$$

12. Show that $u_{m0}^* \equiv 0$ and u_{m0} is identical with (11) in the current section.

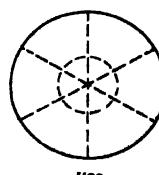
13. Show that u_{mn} has $m + n - 1$ nodal lines.



u_{11}



u_{12}



u_{23}

Fig. 267. Nodal lines of some of the solutions (20).

Graph the nodal lines of the following solutions.

$$14. u_{3n}, n = 1, 2, 3$$

$$15. u_{4n}, n = 1, 2, 3$$

$$16. u_{mn}^*, m, n = 1, 2, 3$$

17. Show that the initial condition $u_t(r, \theta, 0) = 0$ leads to $B_{mn} = 0$, $B_{mn}^* = 0$ in (20).

Find the deflection $u(r, t)$ of the circular membrane of radius $R = 1$, if $c = 1$, the initial velocity is zero, and the initial deflection is $f(r)$ where

$$18. f = 0.1 J_0(\alpha_2 r).$$

$$19. f = k(1 - r^2).$$

$$20. f = k(1 - r^4).$$

Hint: remember Probs. 3–10, Sec. 8.13.

9.11 Laplace's Equation. Potential

One of the most important partial differential equations appearing in physics is Laplace's equation

$$(1) \quad \nabla^2 u = 0.$$

Here $\nabla^2 u$ is the Laplacian of u . With respect to Cartesian coordinates x, y, z in space,

$$(2) \quad \nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}.$$

The theory of the solutions of Laplace's equation is called **potential theory**. Solutions of (1) that have *continuous* second-order partial derivatives are called **harmonic functions**.

The two-dimensional case, when u depends on two variables only, can most conveniently be treated by methods of complex analysis and will be considered in Sec. 10.3 and Chap. 16.

To illustrate the importance of Laplace's equation in engineering mathematics, let us mention some basic applications.

Laplace's equation occurs in connection with gravitational forces. In Ex. 4, Sec. 5.16, we have seen that if a particle A of mass M is fixed at a point (ξ, η, ζ) and another particle B of mass m is at a point (x, y, z) , then A attracts B , the gravitational force being the gradient of the scalar function

$$u(x, y, z) = \frac{c}{r}, \quad c = GMm = \text{const},$$

$$r = \sqrt{(x - \xi)^2 + (y - \eta)^2 + (z - \zeta)^2} \quad (> 0).$$

This function is called the potential of the gravitational field, and it satisfies Laplace's equation.

The extension to the potential and force due to a continuous distribution of mass is quite direct. If a mass of density $\rho(\xi, \eta, \zeta)$ is distributed throughout a region R in space, then the corresponding potential at a point not occupied by mass is defined to be

$$(3) \quad u(x, y, z) = k \iiint_R \frac{\rho}{r} d\xi d\eta d\zeta \quad (k > 0).$$

Since $1/r$ ($r > 0$) is a solution of (1), that is, $\nabla^2(1/r) = 0$, and ρ does not depend on x, y, z , we obtain

$$\nabla^2 u = k \iiint_R \rho \nabla^2 \left(\frac{1}{r}\right) d\xi d\eta d\zeta = 0,$$

that is, the gravitational potential defined by (3) satisfies Laplace's equation at any point which is not occupied by matter.

In *electrostatics* the electrical force of attraction or repulsion between charged particles is governed by *Coulomb's law* (cf. Sec. 5.16), which is of the same mathematical form as Newton's law of gravitation. From this it follows that the field created by a distribution of electrical charges can be described mathematically by a potential function which satisfies Laplace's equation at any point not occupied by charges.

In Chap. 16 we shall see that Laplace's equation also appears in the theory of incompressible fluid flow.

Furthermore, the basic equation in problems of heat conduction is the heat equation

$$u_t = c^2 \nabla^2 u \quad (\text{cf. Sec. 6.9}),$$

and if the temperature u is independent of time t ("steady state"), this equation reduces to Laplace's equation.

In most applications leading to Laplace's equation, it is required to solve a **boundary value problem**, that is, to determine the solution of (1) satisfying given boundary conditions on certain surfaces. It is then necessary to introduce coordinates in space such that those surfaces can be represented in a simple manner. This requires the transformation of the Laplacian (2) into other coordinate systems. Of course, such a transformation is quite similar to that in the case of the Laplacian of a function of two variables.

From (4) in Sec. 9.9 it follows immediately that the Laplacian of a function u in cylindrical coordinates⁸ (cf. Fig. 268 on the next page)

$$(4) \quad r = \sqrt{x^2 + y^2} \quad \theta = \arctan \frac{y}{x}, \quad z = z$$

is

$$(5) \quad \nabla^2 u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} + \frac{\partial^2 u}{\partial z^2}.$$

Other important coordinates are **spherical coordinates**⁹ r, θ, ϕ which are related to Cartesian coordinates as follows (cf. Fig. 269 on the next page):

$$(6) \quad x = r \cos \theta \sin \phi, \quad y = r \sin \theta \sin \phi, \quad z = r \cos \phi.$$

⁸ Observe that θ is not completely determined by the ratio y/x but we must also take the signs of x and y into account.

⁹ Sometimes also called *polar coordinates*.

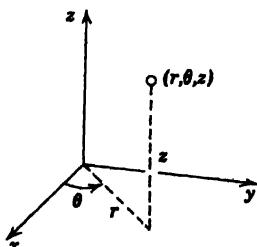


Fig. 268. Cylindrical coordinates.

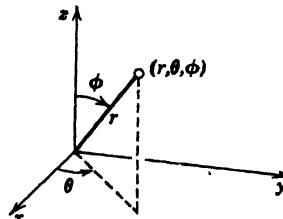


Fig. 269. Spherical coordinates.

The Laplacian of a function u in spherical coordinates is

$$(7) \quad \nabla^2 u = \frac{\partial^2 u}{\partial r^2} + \frac{2}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \phi^2} + \frac{\cot \phi}{r^2} \frac{\partial u}{\partial \phi} + \frac{1}{r^2 \sin^2 \phi} \frac{\partial^2 u}{\partial \theta^2}.$$

This may also be written

$$(7') \quad \nabla^2 u = \frac{1}{r^2} \left[\frac{\partial}{\partial r} \left(r^2 \frac{\partial u}{\partial r} \right) + \frac{1}{\sin \phi} \frac{\partial}{\partial \phi} \left(\sin \phi \frac{\partial u}{\partial \phi} \right) + \frac{1}{\sin^2 \phi} \frac{\partial^2 u}{\partial \theta^2} \right].$$

This formula can be derived in a manner similar to that in Sec. 9.9; the details are left as an exercise for the reader.

Problems

- Find the Laplacian in rectangular coordinates $x^* = ax$, $y^* = by$, $z^* = cz$ where x, y, z are Cartesian coordinates.
- Starting from (2), verify by calculation that with respect to the Cartesian coordinates defined by (9), (10), Sec. 5.17,

$$\nabla^2 u = u_{x^*x^*} + u_{y^*y^*} + u_{z^*z^*}.$$

Show that the following functions $u = f(x, y)$ satisfy the Laplace equation and plot some of the equipotential lines $u = \text{const.}$

- $x^2 - y^2$
- $x^3 - 3xy^2$
- $5. x/(x^2 + y^2)$
- $6. y/(x^2 + y^2)$
- $7. (x^2 - y^2)/(x^2 + y^2)^2$
- Express the spherical coordinates defined by (6) in terms of Cartesian coordinates.
- Verify (5) by transforming $\nabla^2 u$ back into Cartesian coordinates.
- Verify that $u = c/r$ satisfies Laplace's equation in spherical coordinates.
- Show that the only solution of $\nabla^2 u = 0$, depending only on $r = \sqrt{x^2 + y^2 + z^2}$, is $u = c/r + k$; here c and k are constants.
- Determine c and k in Prob. 11 such that u represents the electrostatic potential between two concentric spheres of radii $r_1 = 10$ cm and $r_2 = 2$ cm kept at the potentials $U_1 = 110$ volts and $U_2 = 10$ volts, respectively.
- Show that the only solution of the two-dimensional Laplace equation depending only on $r = \sqrt{x^2 + y^2}$, is $u = c \ln r + k$.
- Find the electrostatic potential between two coaxial cylinders of radii $r_1 = 10$ cm and $r_2 = 2$ cm kept at the potentials $U_1 = 110$ volts and $U_2 = 10$ volts, respectively. Graph and compare the solutions of Probs. 12 and 14.

15. If $u(r, \theta)$ satisfies $\nabla^2 u = 0$, show that $v(r, \theta) = u(r^{-1}, \theta)$ satisfies $\nabla^2 v = 0$. (r and θ are polar coordinates.)
16. Let r, θ, ϕ be spherical coordinates. If $u(r, \theta, \phi)$ satisfies $\nabla^2 u = 0$, show that $v(r, \theta, \phi) = r^{-1}u(r^{-1}, \theta, \phi)$ satisfies $\nabla^2 v = 0$.
17. Show that substitution of $u = U(x, y, z)e^{-i\omega t}$ ($i = \sqrt{-1}$) into the three-dimensional wave equation $u_{tt} = c^2 \nabla^2 u$ yields the so-called *Helmholtz*¹⁰ equation

$$\nabla^2 U + k^2 U = 0 \quad k = \omega/c.$$

Find the steady-state (time-independent) temperature distribution:

18. Between two parallel plates $x = x_0$ and $x = x_1$ kept at the temperatures u_0 and u_1 , respectively.
19. Between two coaxial circular cylinders of radii r_0 and r_1 kept at the temperatures u_0 and u_1 , respectively.
20. Between two concentric spheres of radii r_0 and r_1 kept at the temperatures u_0 and u_1 , respectively.

9.12 Laplace's Equation in Spherical Coordinates. Legendre's Equation

Let us consider a typical boundary value problem that involves Laplace's equation in spherical coordinates. Suppose that a sphere S of radius R is kept at a fixed distribution of electric potential

$$(1) \quad u(R, \theta, \phi) = f(\phi)$$

where r, θ, ϕ are the spherical coordinates defined in the last section, with the origin at the center of S , and $f(\phi)$ is a given function. We wish to find the potential u at all points in space which is assumed to be free of further charges. Since the potential on S is independent of θ , so is the potential in space. Thus $\partial^2 u / \partial \theta^2 = 0$, and from (7') in Sec. 9.11 we see that Laplace's equation reduces to

$$(2) \quad \frac{\partial}{\partial r} \left(r^2 \frac{\partial u}{\partial r} \right) + \frac{1}{\sin \phi} \frac{\partial}{\partial \phi} \left(\sin \phi \frac{\partial u}{\partial \phi} \right) = 0.$$

Furthermore, at infinity the potential will be zero:

$$(3) \quad \lim_{r \rightarrow \infty} u(r, \phi) = 0.$$

We shall solve the boundary value problem consisting of the equation (2) and the boundary conditions (1) and (3) by the method of separating variables. Substituting a solution of the form

$$u(r, \phi) = G(r)H(\phi)$$

¹⁰ HERMANN VON HELMHOLTZ (1821–1894), German physicist.

into (2), and dividing the resulting equation by GH , we obtain

$$\frac{1}{G} \frac{d}{dr} \left(r^2 \frac{dG}{dr} \right) = - \frac{1}{H \sin \phi} \frac{d}{d\phi} \left(\sin \phi \frac{dH}{d\phi} \right).$$

By the usual argument, the two sides of this equation must be equal to a constant, say, k , so that

$$(4) \quad \frac{1}{\sin \phi} \frac{d}{d\phi} \left(\sin \phi \frac{dH}{d\phi} \right) + kH = 0$$

and

$$\frac{1}{G} \frac{d}{dr} \left(r^2 \frac{dG}{dr} \right) = k.$$

The last equation may be written

$$r^2 G'' + 2rG' - kG = 0.$$

This is Cauchy's equation, and from Sec. 2.7 we know that it has solutions of the form $G = r^\alpha$. These solutions will have a particularly simple form if we change our notation and write $n(n+1)$ for k . In fact, then the equation becomes

$$(5) \quad r^2 G'' + 2rG' - n(n+1)G = 0$$

where n is still arbitrary. By substituting $G = r^\alpha$ into (5) we have

$$[\alpha(\alpha-1) + 2\alpha - n(n+1)]r^\alpha = 0.$$

The zeros of the expression in brackets are $\alpha = n$ and $\alpha = -n-1$. Hence we obtain the solutions

$$(6) \quad G_n(r) = r^n \quad \text{and} \quad G_{n+1}(r) = \frac{1}{r^{n+1}}.$$

Introducing $k = n(n+1)$ in (4) and setting

$$\cos \phi = w,$$

we have $\sin^2 \phi = 1 - w^2$ and

$$\frac{d}{d\phi} = \frac{d}{dw} \frac{dw}{d\phi} = -\sin \phi \frac{d}{dw}.$$

Consequently, (4) takes the form

$$(7) \quad \frac{d}{dw} \left[(1 - w^2) \frac{dH}{dw} \right] + n(n+1)H = 0$$

or

$$(7') \quad (1 - w^2) \frac{d^2H}{dw^2} - 2w \frac{dH}{dw} + n(n+1)H = 0.$$

This is Legendre's equation (cf. Sec. 3.3). For integral¹¹ $n = 0, 1, \dots$, the

¹¹ So far, n was arbitrary since k was arbitrary. It can be shown that the restriction of n to real integral values is necessary to make the solution of (7) continuous, together with its derivative of the first order, in the interval $-1 \leq w \leq 1$ or $0 \leq \phi \leq \pi$.

Legendre polynomials

$$H = P_n(w) = P_n(\cos \phi), \quad n = 0, 1, \dots,$$

are solutions of (7). We thus obtain the following two sequences of solutions $u = GH$ of Laplace's equation (2):

$$u_n(r, \phi) = A_n r^n P_n(\cos \phi), \quad u_n^*(r, \phi) = \frac{B_n}{r^{n+1}} P_n(\cos \phi)$$

where $n = 0, 1, \dots$, and A_n and B_n are constants.

To find a solution of (2), valid at points *inside* the sphere and satisfying (1), we consider the series¹²

$$(8) \quad u(r, \phi) = \sum_{n=0}^{\infty} A_n r^n P_n(\cos \phi).$$

For (8) to satisfy (1) we must have

$$(9) \quad u(R, \phi) = \sum_{n=0}^{\infty} A_n R^n P_n(\cos \phi) = f(\phi);$$

that is, (9) must be the generalized Fourier series of $f(\phi)$ in terms of Legendre polynomials. From (8)–(11) in Sec. 8.12 [with $p(x) = 1$] it follows that

$$A_n R^n = \frac{2n+1}{2} \int_{-1}^1 f(w) P_n(w) dw$$

where $\tilde{f}(w)$ denotes $f(\phi)$ as a function of $w = \cos \phi$. Since $dw = -\sin \phi d\phi$, and the limits of integration -1 and 1 correspond to $\phi = \pi$ and $\phi = 0$, respectively, we also have

$$(10) \quad A_n = \frac{2n+1}{2R^n} \int_0^\pi f(\phi) P_n(\cos \phi) \sin \phi d\phi, \quad n = 0, 1, \dots.$$

Thus the series (8) with coefficients (10) is the solution of our problem for points inside the sphere.

To find the solution *exterior* to the sphere, we cannot use the functions $u_n(r, \phi)$ because these functions do not satisfy (3), but we may use the functions $u_n^*(r, \phi)$, which satisfy (3), and proceed as before. This leads to the solution

$$(11) \quad u(r, \phi) = \sum_{n=0}^{\infty} \frac{B_n}{r^{n+1}} P_n(\cos \phi) \quad (r \geq R)$$

with coefficients

$$(12) \quad B_n = \frac{2n+1}{2} R^{n+1} \int_0^\pi f(\phi) P_n(\cos \phi) \sin \phi d\phi.$$

¹²Convergence will not be considered. It can be shown that if $f(\phi)$ and $f'(\phi)$ are piecewise continuous in the interval $0 \leq \phi \leq \pi$, the series (8) having the coefficients (10) can be differentiated termwise twice with respect to r and with respect to ϕ and the resulting series converge and represent $\partial^2 u / \partial r^2$ and $\partial^2 u / \partial \phi^2$, respectively. Hence the series (8) with coefficients (10) is then the solution of our problem inside the sphere.

Problems

1. Verify by substitution that $u_n(r, \phi)$ and $u_n^*(r, \phi)$, $n = 0, 1, 2$ (p. 531) are solutions of (2).
2. Find the surfaces on which the functions u_1, u_2, u_3 are zero.
3. Graph the functions $P_n(\cos \phi)$ for $n = 0, 1, 2$, [cf. (11'), Sec. 3.3].
4. Graph the functions $P_3(\cos \phi)$ and $P_4(\cos \phi)$.

Let r, θ, ϕ be the spherical coordinates used in the text. Find the potential in the interior of the sphere $R = 1$, assuming that there are no charges in the interior and the potential on the surface is $f(\phi)$ where

5. $f(\phi) = 1$
6. $f(\phi) = \cos \phi$
7. $f(\phi) = \cos^2 \phi$
8. $f(\phi) = \cos^3 \phi$
9. $f(\phi) = \cos 2\phi$
10. $f(\phi) = \cos 3\phi$
11. $f(\phi) = 10 \cos^3 \phi - 3 \cos^2 \phi - 5 \cos \phi - 1$
12. Show that in Prob. 5, the potential exterior to the sphere is the same as that of a point charge at the origin.
13. Graph the intersections of the equipotential surfaces in Prob. 6 with the xz -plane.
14. Find the potential exterior to the sphere in Probs. 5–11.
15. Verify by substitution that the solution of Prob. 7 satisfies (2).
16. Consider a long cable or telephone wire (Fig. 270) which is imperfectly insulated so that leaks occur along the entire length of the cable. The source S of the current $i(x, t)$ in the cable is at $x = 0$, the receiving end T at $x = l$. The current flows

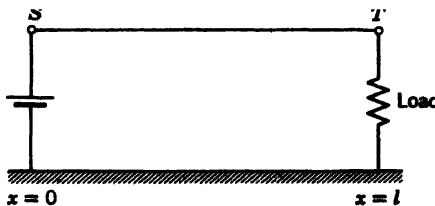


Fig. 270. Transmission line.

from S to T , through the load, and returns to the ground. Let the constants R , L , C , and G denote the resistance, inductance, capacitance to ground and conductance to ground, respectively, of the cable per unit length. Show that

$$-\frac{\partial u}{\partial x} = Ri + L \frac{\partial i}{\partial t} \quad (\text{First transmission line equation})$$

where $u(x, t)$ is the potential in the cable. Hint: apply Kirchhoff's second law to a small portion of the cable between x and $x + \Delta x$ (difference of the potentials at x and $x + \Delta x$ = resistive drop + inductive drop).

17. Show that for the cable in Prob. 16,

$$-\frac{\partial i}{\partial x} = Gu + C \frac{\partial u}{\partial t} \quad (\text{Second transmission line equation}).$$

Hint: use Kirchhoff's first law (difference of the currents at x and $x + \Delta x$ = loss due to leakage to ground + capacitive loss).

18. Show that elimination of i from the transmission line equations leads to

$$u_{xx} = LC u_{tt} + (RC + GL)u_t + RGu.$$

19. For a submarine cable G is negligible and the frequencies are low. Show that this leads to the so-called *submarine cable equations* or *telegraph equations*

$$u_{xx} = RC u_t, \quad i_{xx} = RC i_t.$$

20. Show that in the case of alternating currents of high frequencies the equation in Prob. 18 and the analogous equation for the current i can be approximated by the so-called *high frequency line equations*

$$u_{xx} = LC u_{tt}, \quad i_{xx} = LC i_{tt}.$$

Chapter 10

Complex Analytic Functions

Many engineering problems may be treated and solved by methods of complex analysis. Roughly speaking, these problems can be subdivided into two large classes. The first class consists of "elementary problems" for which the knowledge of complex numbers gained in college algebra and calculus is sufficient. For example, many applications in connection with electric circuits and mechanical vibrating systems are of this type. The second class of problems requires a detailed knowledge of the theory of complex analytic functions ("theory of functions," for short) and the powerful and elegant methods used in this branch of mathematics. Interesting problems in the theory of heat, in fluid dynamics, and in electrostatics belong to this category.

The following chapters will be devoted to the major parts of the theory of complex analytic functions and their applications. We shall see that the importance of these functions in engineering mathematics has the following three main roots.

1. *The real and imaginary parts of an analytic function are solutions of Laplace's equation in two independent variables. Consequently, two-dimensional potential problems can be treated by methods developed in connection with analytic functions.*

2. *Many complicated real and complex integrals that occur in applications can be evaluated by methods of complex integration.*

3. *The majority of nonelementary functions appearing in engineering mathematics are analytic functions, and the consideration of these functions for complex values of the independent variable leads to a much deeper and more detailed knowledge of their properties.*

In the present chapter we shall define analytic functions and consider their fundamental properties. The second half of the chapter will be devoted to the most important elementary complex functions.

Prerequisites for this chapter: elementary calculus, complex numbers (Secs. 0.4 and 0.5).

References: Appendix 1, Part G.

Answers to problems: Appendix 2.

10.1 Complex Numbers. Triangle Inequality

Complex numbers were introduced in Secs. 0.4 and 0.5, and we shall now consider several further properties.

The **real** and **imaginary parts** of a complex number

$$z = x + iy \quad (x, y, \text{ real})$$

are x and y , respectively, and we write

$$\operatorname{Re} z = x, \quad \operatorname{Im} z = y.$$

Introducing polar coordinates r and θ defined by

$$x = r \cos \theta, \quad y = r \sin \theta,$$

we may represent the complex number z in polar form

$$z = r(\cos \theta + i \sin \theta).$$

$|z|$ is called the **absolute value**, and θ the **argument** of z , and we write

$$|z| = r \quad \text{and} \quad \arg z = \theta.$$

Geometrically, $|z|$ is the distance of the point z from the origin (Fig. 271). Hence the inequality

$$|z_1| > |z_2|$$

means that the point z_1 is farther from the origin than the point z_2 , and $|z_1 - z_2|$ is the distance between the points z_1 and z_2 (Fig. 272).

For given z ($\neq 0$), the argument θ is determined only up to integral multiples of 2π . The value of θ which lies in the interval

$$-\pi < \theta \leq \pi$$

is called the **principal value** of the argument of z .

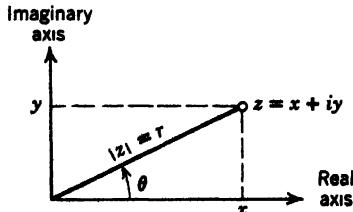


Fig. 271. Complex plane.

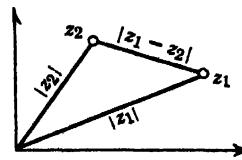


Fig. 272. Distance between two points in the complex plane.

Example 1. Let $z = 1 + i$. Then $z = \sqrt{2} \left(\cos \frac{\pi}{4} + i \sin \frac{\pi}{4} \right)$ and

$$|z| = \sqrt{2}, \quad \arg z = \frac{\pi}{4} \pm 2n\pi \quad (n = 0, 1, \dots).$$

The principal value of $\arg z$ is $\pi/4$ (Fig. 273).

Example 2. Let $z = 3 + 3\sqrt{3}i$. Then $z = 6 \left(\cos \frac{\pi}{3} + i \sin \frac{\pi}{3} \right)$, $|z| = 6$, and the principal value of $\arg z$ is $\pi/3$.

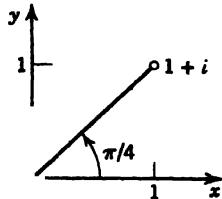


Fig. 273. Example 1.

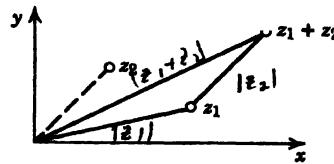


Fig. 274. Triangle inequality.

Since for any complex number $z = x + iy$,

$$|z| = \sqrt{x^2 + y^2} \geq |x| \quad \text{and similarly} \quad |z| \geq |y|,$$

we obtain the useful inequalities

$$(1) \quad |\operatorname{Re} z| \leq |z|, \quad |\operatorname{Im} z| \leq |z|.$$

Let z_1 and z_2 be any complex numbers. Then the origin and the points z_1 and $z_1 + z_2$ are the vertices of a triangle¹ (Fig. 274) whose sides are $|z_1|$, $|z_2|$, and $|z_1 + z_2|$. We thus obtain the important triangle inequality

$$(2) \quad |z_1 + z_2| \leq |z_1| + |z_2|,$$

which will be frequently used in our considerations. The formal proof is left to the student (Prob. 21).

By induction the triangle inequality can be extended to arbitrary sums:

$$(3) \quad |z_1 + z_2 + \cdots + z_n| \leq |z_1| + |z_2| + \cdots + |z_n|;$$

that is, the absolute value of a sum is at most equal to the sum of the absolute values of the summands.

We shall now consider some important types of curves and regions in the complex plane and their representations by equations and inequalities.

Since the distance between two points z and a is $|z - a|$, it follows that a circle C of radius ρ with center at a point a (Fig. 275) can be represented in the form

$$(4) \quad |z - a| = \rho.$$

¹ Which may degenerate if z_1 and z_2 lie on the same straight line through the origin.

Consequently, the inequality

$$(5) \quad |z - a| < \rho$$

holds for any point z inside C ; that is, (5) represents the interior of C . Such a region is called a circular disk, or, more precisely, an open circular disk, in contrast to the closed circular disk

$$|z - a| \leq \rho,$$

which consists of the interior of C and C itself. The open circular disk (5) is also called a neighborhood of the point a . Obviously, a has infinitely many such neighborhoods, each of which corresponds to a certain value of $\rho (> 0)$.

Similarly, the inequality

$$|z - a| > \rho$$

represents the exterior of the circle C . Furthermore, the region between two concentric circles of radii ρ_1 and $\rho_2 (> \rho_1)$ can be represented in the form

$$\rho_1 < |z - a| < \rho_2,$$

where a is the center of the circles. Such a region is called an open circular ring or open annulus (Fig. 276).

The equation

$$|z| = 1$$

represents the so-called unit circle, that is, the circle of radius 1 with center at the origin (Fig. 277). This circle will play an important role in various considerations.

It is important that the student becomes completely familiar with representations of curves and regions in the complex plane. Therefore, he should pay particular attention to the problems included in the present section. This will make his future work less difficult.

We finally define some notions which are of general importance and will be used in our further consideration.

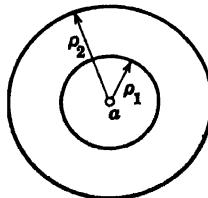


Fig. 276. Annulus in the complex plane.

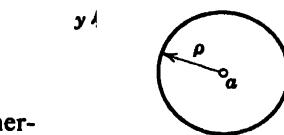


Fig. 275. Circle in the complex plane.

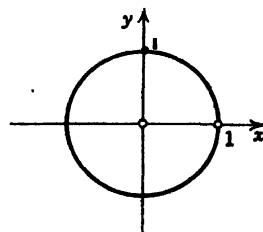


Fig. 277. Unit circle.

The term *set of points* in the complex plane means any sort of collection of finitely or infinitely many points. For example, the solutions of a quadratic equation, the points on a line, and the points in the interior of a circle are sets.

A set S is called *open* if every point of S has a neighborhood every point of which belongs to S . For example, the points in the interior of a circle or a square form an open set, and so do the points of the "right half-plane" $\operatorname{Re} z = x > 0$. The set consisting of the points in the interior of a circle C and on C ("closed circular disk") is not open, because no neighborhood of a point of C lies entirely in the set.

A set S in the complex plane is called *closed* if the points of the plane which do not belong to S form an open set. For example, the points on and inside the unit circle form a closed set.

A set is called *bounded* if all of its points lie within a circle of sufficiently large radius. For example, the points in a rectangle form a bounded set, while the points on a straight line do not form a bounded set.

An open set S is said to be *connected* if any two of its points can be joined by a broken line of finitely many linear segments all of whose points belong to S . An open connected set is called a **domain**. Thus the interior of a circle is a domain.

A *boundary point* of a set S is a point every neighborhood of which contains both points which belong to S and points which do not belong to S . For example, the boundary points of an annulus are the points on the two bounding circles. Clearly, if a set S is open, then no boundary point belongs to S ; if S is closed, then every boundary point belongs to S .

A **region** is a set consisting of a domain plus, perhaps, some or all of its boundary points. (We should note that some authors use the term "region" for what we call a domain, and others make no distinction between the two terms.)

Problems

Find

1. $\operatorname{Re} \frac{1}{2+i}$

2. $\operatorname{Im} \frac{2+i}{3+4i}$

3. $\operatorname{Re} z^3$

4. $\operatorname{Im} z^4$

5. $\operatorname{Re} \frac{(1+i)^2}{3+2i}$

6. $\operatorname{Im} \frac{2-i}{4-3i}$

7. $\operatorname{Im} \frac{1}{z^2}$

8. $|z|^2$

9. $\left| \frac{1+4i}{4+i} \right|$

10. $\left| \frac{z-1}{z+1} \right|$

11. $|\cos \theta + i \sin \theta|$

Determine the principal value of the arguments of:

12. -3

13. $3+3i$

14. $1-i\sqrt{3}$

15. $-4i$

16. Show that the conjugate of a sum (product) of complex numbers is the sum (product) of the conjugates of those numbers.

17. Show that multiplication of a complex number by i corresponds to a counterclockwise rotation of the corresponding vector through the angle $\pi/2$.

18. Verify (1) for $z = (4 + 5i)^2$ and $z = (-5 + 2i)^3$.

Verify the triangle inequality for

19. $z_1 = 1 + i, z_2 = -2 + 3i$

20. $z_1 = 5 + 2i, z_2 = 1 - 4i$

21. Prove the triangle inequality.

22. Under what conditions will $|z_1 + z_2| = |z_1| + |z_2|$?

23. Using the triangle inequality, prove that $|z_1 + z_2| \geq |z_1| - |z_2|$.

24. Using the triangle inequality, prove that $|z_1 + z_2| \geq |z_2| - |z_1|$.

25. Show that $(|x| + |y|)/\sqrt{2} \leq |z| \leq |x| + |y|$. Give examples.

26. Show that $|z_1 + z_2|^2 + |z_1 - z_2|^2 = 2(|z_1|^2 + |z_2|^2)$.

Determine and graph the loci represented by

27. $\operatorname{Im} z \geq -1$

28. $\operatorname{Im}(z^2) \leq 2$

29. $\operatorname{Re}(z^2) \geq 1$

30. $|\arg z| < \pi/4$

31. $-\pi < \operatorname{Re} z < \pi$

32. $|1/z| < 1$

33. $\left| \frac{z+1}{z-1} \right| = 2$

34. $\left| \frac{z+i}{z-i} \right| = 1$

35. $\left| \frac{z+1}{z-1} \right| = 1$

10.2 Limit. Derivative. Analytic Function

We shall now define some basic concepts of complex analysis. The reader will notice that these definitions are quite similar to those in real calculus.

If x and y are real variables, then $z = x + iy$ is said to be a *complex variable*.

Consider now two complex variables z and w , and suppose that a relation is given so that to each value in some region of the complex z -plane there corresponds one or more values of w in a well-defined manner. Then w is said to be a *function* of z , defined in that region, and we write

$$w = f(z)$$

or $w = g(z)$, etc., or sometimes simply $w(z)$. The set of complex numbers which $w = f(z)$ can assume as z varies in that region, is called the *range of values* of the function $w = f(z)$.

If to each z in that region there corresponds only one value of $w = f(z)$, then the function $f(z)$ is said to be *single-valued*. A function which is not single-valued, is called *multi-valued*.

General assumption. Unless it is stated otherwise, we shall assume that the functions under consideration are single-valued.

Let u and v be the real and imaginary parts of w . Then, since w depends on $z = x + iy$, it is clear that in general u depends on x and y , and so does v . We may, therefore, write

$$w = f(z) = u(x, y) + iv(x, y),$$

•

and this shows that a complex function $f(z)$ is equivalent to two real functions $u(x, y)$ and $v(x, y)$.

Example 1. Let

$$w = f(z) = z^2 + 3z.$$

Then

$$u(x, y) = \operatorname{Re} f(z) = x^2 - y^2 + 3x \quad \text{and} \quad v(x, y) = \operatorname{Im} f(z) = 2xy + 3y.$$

At $z = x + iy = 1 + 3i$ the function has the value

$$(1 + 3i)^2 + 3(1 + 3i) = -5 + 15i,$$

and we may write

$$f(1 + 3i) = -5 + 15i, \quad u(1, 3) = -5, \quad v(1, 3) = 15.$$

Similarly, $f(1 + i) = 3 + 5i$, etc. Obviously, our function is defined for all z .

Example 2. Determine the value of $f(z) = 3z = 3x - 3iy$ at $z = 2 + 4i$. Since $x = 2$ and $y = 4$, we obtain $f(2 + 4i) = 6 - 12i$.

A function $f(z)$ is said to have the limit l as z approaches z_0 if $f(z)$ is defined in a neighborhood of z_0 (except perhaps at z_0) and if for every positive real number ϵ (no matter how small but not zero) we can find a positive real number δ such that

$$(1) \quad |f(z) - l| < \epsilon$$

for all values $z \neq z_0$ in the disk $|z - z_0| < \delta$. This means that the values of $f(z)$ are as close as desired to l for all z which are sufficiently close to z_0 (Fig. 278), and we write

$$(2) \quad \lim_{z \rightarrow z_0} f(z) = l.$$

Note well that this definition of a limit implies that z may approach z_0 from any direction in the complex plane.

The function $f(z)$ is said to be continuous at $z = z_0$ if $f(z_0)$ is defined and

$$(3) \quad \lim_{z \rightarrow z_0} f(z) = f(z_0).$$

Note that by the definition of a limit this implies that $f(z)$ is defined in some neighborhood of z_0 . A *continuous function*, without further qualification, is one which is continuous at all points where it is defined.

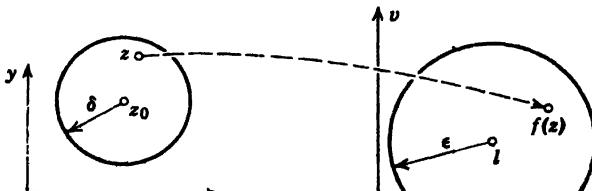


Fig. 278. Limit.

A function $f(z)$ is said to be *differentiable* at a point $z = z_0$ if the limit

$$(4) \quad f'(z_0) = \lim_{\Delta z \rightarrow 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z}$$

exists. This limit is then called the *derivative* of $f(z)$ at the point $z = z_0$.

Setting $z_0 + \Delta z = z$, we have $\Delta z = z - z_0$ and may also write

$$(4') \quad f'(z_0) = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}.$$

Since the notion of a limit implies that z may approach z_0 from any direction, differentiability means that, along whatever path z approaches z_0 , the quotient in (4') always approaches a certain value and all these values are equal. This fact will be quite important in our later considerations.

By definition of a limit, (4') means that there is a complex number $f'(z_0)$ for which, an $\epsilon > 0$ being given, we can find a $\delta > 0$ such that

$$(5) \quad \left| \frac{f(z) - f(z_0)}{z - z_0} - f'(z_0) \right| < \epsilon \quad \text{when} \quad |z - z_0| < \delta.$$

Example 3. The derivative of the function $f(z) = z^2$ is $2z$ at any point z , because

$$f'(z) = \lim_{\Delta z \rightarrow 0} \frac{(z + \Delta z)^2 - z^2}{\Delta z} = \lim_{\Delta z \rightarrow 0} (2z + \Delta z) = 2z.$$

All the familiar rules of real differential calculus, such as the rules for differentiating a constant, integer powers of z , sums, products, and quotients of differentiable functions, and the chain rule for differentiating a function of a function, continue to hold in complex.

In fact, the corresponding proofs are literally the same.

It is important to note that there are many simple functions which do not have a derivative at any point. For instance, $f(z) = \bar{z} = x - iy$ is such a function, because, setting $\Delta z = \Delta x + i \Delta y$, we see that, as Δz approaches zero, the quotient

$$(6) \quad \frac{f(z + \Delta z) - f(z)}{\Delta z} = \frac{[(x + \Delta x) - i(y + \Delta y)] - (x - iy)}{\Delta x + i \Delta y} = \frac{\Delta x - i \Delta y}{\Delta x + i \Delta y}$$

approaches a value which depends on the manner in which Δx and Δy approach zero. In fact, if we let $\Delta y \rightarrow 0$ first, and then $\Delta x \rightarrow 0$ (path I in Fig. 279), the quotient approaches $+1$, but if $\Delta x \rightarrow 0$ first, and then $\Delta y \rightarrow 0$ (path II in Fig. 279), the quotient approaches -1 . This example may be surprising, but it merely illustrates that differentiability of a complex function is a rather severe requirement.

We are now in a position to state the following basic definition.

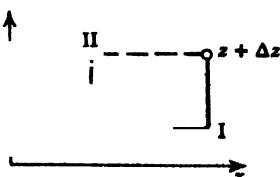


Fig. 279. Paths in (6).

Definition (Analyticity.) A function $f(z)$ is said to be analytic at a point $z = z_0$ if it is defined, and has a derivative, at every point in some neighborhood of z_0 . It is said to be analytic in a domain D if it is analytic at every point in D .

We shall also use the term **analytic function**, without reference to a specific domain; by this term we shall mean a function which is analytic in *some* domain D .

Example 4. The integral powers $1, z, z^2, \dots$ and, more generally, **polynomials**, that is, functions of the form

$$f(z) = c_0 + c_1 z + c_2 z^2 + \dots + c_n z^n$$

where c_0, \dots, c_n are complex constants, are analytic in the entire plane. The function $1/(1 - z)$ is analytic everywhere except at $z = 1$.

Instead of *analytic in D* the terms *holomorphic in D* and *regular in D* are also used in the literature.

Complex analysis is concerned exclusively with analytic functions, and although many simple functions are not analytic, the large variety of remaining functions will yield a branch of mathematics which is most beautiful from the theoretical point of view and most useful for practical purposes.

Problems

Find $f(2 + i), f(3i), f(-4 + i)$ where $f(z)$ equals

$$1. 3z^2 + z \quad 2. 1/z^2 \quad 3. (z + 1)/(z - 1)$$

Find the real and imaginary parts of the following functions.

$$4. f(z) = 2z^2 - 3z \quad 5. f(z) = 1/(1 - z) \quad 6. f(z) = z^3 - z^2 + z$$

Suppose that z varies in a region R in the z -plane. Find the (precise) region in the w -plane in which the corresponding values of $w = f(z)$ lie, and show the two regions graphically.

$$\begin{array}{ll} 7. f(z) = 3z, |\arg z| < \pi/2 & 8. f(z) = 1/z, \operatorname{Re} z > 0 \\ 9. f(z) = z^2, |z| > 3 & 10. f(z) = z^3, |\arg z| \leq \pi/4 \\ 11. f(z) = 1/z^2, |\arg z| \leq \pi/4 & 12. f(z) = z^2 + 1, |z| \leq 2 \end{array}$$

13. Prove that (2) is equivalent to the pair of relations

$$\lim_{z \rightarrow z_0} \operatorname{Re} f(z) = \operatorname{Re} l, \quad \lim_{z \rightarrow z_0} \operatorname{Im} f(z) = \operatorname{Im} l.$$

14. If $\lim_{z \rightarrow z_0} f(z) = l$ and $\lim_{z \rightarrow z_0} g(z) = p$, show that

$$\begin{aligned} \lim_{z \rightarrow z_0} [f(z) + g(z)] &= \lim_{z \rightarrow z_0} f(z) + \lim_{z \rightarrow z_0} g(z) = l + p, \\ \lim_{z \rightarrow z_0} [f(z)g(z)] &= \lim_{z \rightarrow z_0} f(z) \lim_{z \rightarrow z_0} g(z) = lp. \end{aligned}$$

Are the following functions continuous at the origin?

15. $f(z) = \operatorname{Re} z/|z|$ when $z \neq 0, f(0) = 0$
16. $f(z) = \operatorname{Im} z/(1 + |z|)$ when $z \neq 0, f(0) = 0$
17. $f(z) = (\operatorname{Re} z)^2/|z|$ when $z \neq 0, f(0) = 0$
18. $f(z) = \operatorname{Re}(z^3)/|z^3|$ when $z \neq 0, f(0) = 0$

19. If z_1, z_2, \dots are complex numbers for which $\lim_{n \rightarrow \infty} z_n = a$ and $f(z)$ is continuous at $z = a$, show that $\lim_{n \rightarrow \infty} f(z_n) = f(a)$.
20. Using the definition of limit, show that $f(z) = z^2$ is continuous.

Show that

$$21. [af(z) + bg(z)]' = af'(z) + bg'(z) \quad 22. [f(z)g(z)]' = f'(z)g(z) + f(z)g'(z)$$

Differentiate the following functions.

$$23. (z^2 + 1)^2$$

$$24. 1/(1 - z)$$

$$25. (z + 1)^2/(z^2 + 1)$$

Find the value of the derivative at z_0 :

$$26. f(z) = z^3 - 2z, \quad z_0 = 1 - i$$

$$27. f(z) = (z + 2i)/(z - 2i), \quad z_0 = 3 + i$$

$$28. f(z) = (z^2 - 1)^2, \quad z_0 = i$$

$$29. f(z) = iz^2 + (1 + i)z, \quad z_0 = -2 + i$$

30. Show that $f(z) = |z|^2$ is differentiable only at $z = 0$. Hint: use the relation $|z + \Delta z|^2 = (z + \Delta z)(\bar{z} + \bar{\Delta z})$.

✓10.3 Cauchy-Riemann Equations. Laplace's Equation

We shall now derive a simple basic criterion for analyticity of a complex function

$$(1) \quad w = f(z) = u(x, y) + iv(x, y).$$

We first suppose that $f(z)$ is analytic in a domain D of the z -plane. Then, by definition, $f(z)$ has a derivative

$$(2) \quad f'(z) = \lim_{\Delta z \rightarrow 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}$$

everywhere in D , where Δz may approach 0 along any path. We may set $\Delta z = \Delta x + i \Delta y$. Choosing path I in Fig. 280, we let $\Delta y \rightarrow 0$ first and then $\Delta x \rightarrow 0$. After Δy becomes zero, $\Delta z = \Delta x$, and by (1),

$$\begin{aligned} f'(z) &= \lim_{\Delta x \rightarrow 0} \frac{u(x + \Delta x, y) + iv(x + \Delta x, y) - [u(x, y) + iv(x, y)]}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{u(x + \Delta x, y) - u(x, y)}{\Delta x} + i \lim_{\Delta x \rightarrow 0} \frac{v(x + \Delta x, y) - v(x, y)}{\Delta x} \end{aligned}$$

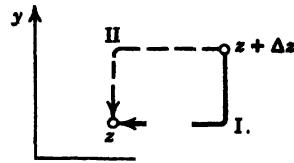


Fig. 280. Paths in (2).

Since $f'(z)$ exists, the last two real limits exist. They are the partial derivatives of u and v with respect to x . Hence $f'(z)$ can be written

$$(3) \quad f'(z) = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x}.$$

Similarly, if we choose path II in Fig. 280, we let $\Delta x \rightarrow 0$ first and then $\Delta y \rightarrow 0$. After Δx becomes zero, $\Delta z = i \Delta y$, and

$$f'(z) = \lim_{\Delta y \rightarrow 0} \frac{u(x, y + \Delta y) - u(x, y)}{i \Delta y} + i \lim_{\Delta y \rightarrow 0} \frac{v(x, y + \Delta y) - v(x, y)}{i \Delta y};$$

that is,

$$(4) \quad f'(z) = -i \frac{\partial u}{\partial y} + \frac{\partial v}{\partial y}$$

because $1/i = -i$. The existence of $f'(z)$ thus implies the existence of the four partial derivatives in (3) and (4).

By equating the real and the imaginary parts of the right-hand sides of (3) and (4) we obtain

$$(5) \quad \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \text{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

These basic relations are called the **Cauchy-Riemann differential equations**.^{2,3}

We may sum up our result as follows.

Theorem 1. *The real part and the imaginary part of an analytic function $f(z) = u(x, y) + i v(x, y)$ satisfy the Cauchy-Riemann equations (5) at each point where $f(z)$ is analytic.*

Example 1. The function $f(z) = z^2 = x^2 - y^2 + 2ixy$ is analytic for all z , and $f'(z) = 2z$. We have $u = x^2 - y^2$, $v = 2xy$,

$$\frac{\partial u}{\partial x} = 2x, \quad \frac{\partial u}{\partial y} = -2y, \quad \frac{\partial v}{\partial x} = 2y, \quad \frac{\partial v}{\partial y} = 2x,$$

and the Cauchy-Riemann equations are satisfied for all x, y .

The Cauchy-Riemann equations are fundamental because they are not only necessary but also sufficient for a function to be analytic; more precisely:

Theorem 2. *If two real (single-valued) functions $u(x, y)$ and $v(x, y)$ of two real variables x and y have continuous first partial derivatives that satisfy the Cauchy-Riemann equations in some domain D , then the complex function $f(z) = u(x, y) + i v(x, y)$ is analytic in D .*

Proof. Let $P: (x, y)$ be any fixed point in D . Since D is a domain, it contains a neighborhood of P . In this neighborhood we choose a point

² Cf. the footnote on p. 87.

³ BERNHARD RIEMANN (1826–1866), German mathematician, who developed what may be called the “geometrical approach” to complex analysis, based upon the Cauchy-Riemann equations and conformal mapping, in contrast to the German mathematician KARL WEIERSTRASS (1815–1897), who based complex analysis on power series (cf. Sec. 13.6, 14.1). Riemann developed also the so-called Riemannian geometry, which is the mathematical base of Einstein’s theory of relativity.

$Q: (x + \Delta x, y + \Delta y)$. Then the segment PQ is in D , and from the mean-value theorem in Sec. 5.15 it follows that

$$(6) \quad \begin{aligned} u(x + \Delta x, y + \Delta y) - u(x, y) &= \Delta x \frac{\partial u}{\partial x} + \Delta y \frac{\partial u}{\partial y}, \\ v(x + \Delta x, y + \Delta y) - v(x, y) &= \Delta x \frac{\partial v}{\partial x} + \Delta y \frac{\partial v}{\partial y}, \end{aligned}$$

the derivatives being evaluated at a suitable point of that segment. We set

$$f(z) = u(x, y) + iv(x, y) \quad \text{and} \quad \Delta z = \Delta x + i \Delta y.$$

Then we obtain from (6)

$$f(z + \Delta z) - f(z) = \frac{\partial u}{\partial x} \Delta x + \frac{\partial u}{\partial y} \Delta y + i \left(\frac{\partial v}{\partial x} \Delta x + \frac{\partial v}{\partial y} \Delta y \right).$$

Using the Cauchy-Riemann equations, we may replace $\partial u / \partial y$ by $-\partial v / \partial x$ and $\partial v / \partial y$ by $\partial u / \partial x$, finding

$$f(z + \Delta z) - f(z) = \frac{\partial u}{\partial x} (\Delta x + i \Delta y) + i \frac{\partial v}{\partial x} (\Delta x + i \Delta y) = \left(\frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \right) \Delta z.$$

Dividing both sides by Δz , we obtain

$$\frac{f(z + \Delta z) - f(z)}{\Delta z} = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x}.$$

If we let Δz approach zero, then the derivatives on the right become $\partial u / \partial x$ and $\partial v / \partial x$, evaluated at (x, y) ; hence the limit of the quotient on the left exists, and from (3) we see that it is $f'(z)$. This means that $f(z)$ is analytic in D , and the proof is complete.

These theorems are of great practical importance, and we may now easily decide whether or not a given complex function is analytic.

Example 2. Let

$$f(z) = \operatorname{Re} z = x.$$

Then

$$u = x \quad \text{and} \quad v = 0.$$

The equations (5) are not satisfied, that is, $\operatorname{Re} z$ is not analytic. Similarly, $\operatorname{Im} z$ is not analytic. Further simple functions which are not analytic are included in the problems.

We mention that if we use the polar form $z = r(\cos \theta + i \sin \theta)$ and set

$$f(z) = u(r, \theta) + iv(r, \theta),$$

then the Cauchy-Riemann equations are

$$(7) \quad \frac{\partial u}{\partial r} = \frac{1}{r} \frac{\partial v}{\partial \theta}, \quad \frac{\partial v}{\partial r} = -\frac{1}{r} \frac{\partial u}{\partial \theta}.$$

We shall prove later (in Sec. 12.6) that the derivative of an analytic function $f(z) = u(x, y) + iv(x, y)$ is itself analytic. By this important fact $u(x, y)$ and $v(x, y)$ will have continuous partial derivatives of all orders. In particular, the mixed second derivatives of these functions will be equal:

$$\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial y \partial x}, \quad \frac{\partial^2 v}{\partial x \partial y} = \frac{\partial^2 v}{\partial y \partial x}.$$

Differentiating the Cauchy-Riemann equations, we thus obtain

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 v}{\partial x \partial y}, \quad \frac{\partial^2 u}{\partial y^2} = -\frac{\partial^2 v}{\partial x \partial y}$$

$$\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 v}{\partial y^2}, \quad \frac{\partial^2 u}{\partial x \partial y} = -\frac{\partial^2 v}{\partial x^2}.$$

This yields the following important result.

Theorem 3. *The real part and the imaginary part of a complex function $f(z) = u(x, y) + iv(x, y)$ that is analytic in a domain D are solutions of Laplace's equation,*

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad \nabla^2 v = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0,$$

in D and have continuous second partial derivatives in D .

This is one of the main reasons for the great practical importance of complex analysis in engineering mathematics, as we shall see in Chaps. 11 and 16.

A solution of Laplace's equation having *continuous* second-order partial derivatives is called a **harmonic function** (cf. also Sec. 9.11). Hence the real and imaginary parts of an analytic function are harmonic functions.

If two harmonic functions $u(x, y)$ and $v(x, y)$ satisfy the Cauchy-Riemann equations in a domain D , that is, if u and v are the real and imaginary parts of an analytic function $f(z)$ in D , then $v(x, y)$ is said to be a **conjugate harmonic function** of $u(x, y)$ in D , and the pair u, v is called a *pair of conjugate harmonic functions* in D . (Of course this use of the word "conjugate" is different from that employed in defining \bar{z} , the conjugate of a complex number z .)

A conjugate of a given harmonic function can be obtained from the Cauchy-Riemann equations, as may be illustrated by the following example.

Example 3. The function $u = x^2 - y^2$ is harmonic, and we have $\partial u / \partial x = 2x$, $\partial u / \partial y = -2y$. Hence a conjugate of u must satisfy

$$\frac{\partial v}{\partial y} = 2x, \quad \frac{\partial v}{\partial x} = 2y.$$

By integrating the first equation with respect to y we obtain

$$v = 2xy + h(x)$$

where $h(x)$ depends only on x . Substituting this in the last equation, we have $h'(x) = 0$ and, therefore, $h = c = \text{const}$. Hence, the most general conjugate function of $x^2 - y^2$ is $2xy + c$, where c is a real constant, and the most general analytic function with the real part $x^2 - y^2$ is $x^2 - y^2 + i(2xy + c) = z^2 + ic$.

Problems

Are the following functions analytic?

1. $f(z) = z^3 + z$
2. $f(z) = \operatorname{Im} z$
3. $f(z) = z$
4. $f(z) = |z|^2$
5. $f(z) = 1/(1 - z)$
6. $f(z) = e^x(\cos y + i \sin y)$
7. $f(z) = 1/z^2$
8. $f(z) = \sin x \cosh y + i \cos x \sinh y$
9. $f(z) = \arg z$
10. $f(z) = \cos x \cosh y - i \sin x \sinh y$

Using (3) or (4), find the derivative of the following functions.

11. $f(z) = az + b$ 12. $f(z) = z^2$ 13. $f(z) = 1/z$

14. Show that

$$(8) \quad f'(z) = \frac{\partial u}{\partial x} - i \frac{\partial u}{\partial y}, \quad f'(z) = \frac{\partial v}{\partial y} + i \frac{\partial v}{\partial x}.$$

15. If $f(z)$ is analytic in a domain D and $|f(z)| = \text{const}$ in D , show that $f(z) = \text{const}$.

16. If $f(z)$ is analytic in a domain D and $\operatorname{Re} f(z) = \text{const}$ in D , show that $f = \text{const}$.

17. If $f(z)$ is analytic in a domain D and $f'(z) = 0$ everywhere in D , show that $f(z) = \text{const}$.

18. Derive (7) from (5).

Using (7), show that the following functions are analytic.

19. $f(z) = z^4$ 20. $f(z) = \ln r + i\theta$ 21. $f(z) = 1/z^3$ ($z \neq 0$)

Show that the following functions are harmonic and find a corresponding analytic function $f(z) = u(x, y) + iv(x, y)$.

22. $u = x$	23. $v = xy$	24. $u = xy$
25. $u = e^x \cos y$	26. $u = \ln(x^2 + y^2)$	27. $v = \arg z$
28. $u = \sin x \cosh y$	29. $v = -\sin x \sinh y$	30. $u = x/(x^2 + y^2)$

10.4 Rational Functions. Root

The remaining sections of this chapter will be devoted to the most important elementary complex functions, such as powers, exponential function, logarithm, trigonometric functions, etc. We shall see that these functions can easily be defined in such a way that, for real values of the independent variable, the functions become identical with the familiar real functions. Some of the complex functions have interesting properties, which do not show when the independent variable is restricted to real values. The student should follow the consideration with great care, because these elementary functions will frequently be needed in applications. Furthermore, a detailed knowledge of these special functions will be helpful later for a better understanding of our more general considerations.

The powers

(1) $w = z^n$ $n = 0, 1, \dots$

are analytic in the entire plane, and the same is true for a function of the form

(2) $w = c_0 + c_1z + c_2z^2 + \dots + c_nz^n$ ($c_n \neq 0$)

where c_0, \dots, c_n are (complex or real) constants. Such a function is called a **polynomial** or an *entire rational function*. The exponent n is called the *degree* of the polynomial. The study of these functions is the principal subject of classical algebra.

A quotient of two polynomials $p(z)$ and $q(z)$ is called a (*fractional*) **rational function**. Such a function

(3) $w = \frac{p(z)}{q(z)}$

is analytic for every z for which $q(z)$ is not zero. A rational function of the particularly simple form

$$\frac{c}{(z - z_0)^m} \quad (c \neq 0)$$

where both c and z_0 are complex numbers and m is a positive integer, is called a *partial fraction*. It is proved in algebra that every rational function can be represented as the sum of a polynomial and finitely many partial fractions.

If $z = w^n$ ($n = 1, 2, \dots$), then to each value of w there corresponds one value of z . We shall immediately see that to a given $z \neq 0$ there correspond precisely n distinct values of w . Each of these values is called an *n th root* of z , and we write

(4) $w = \sqrt[n]{z}$.

Hence this symbol is *multi-valued*, namely, *n -valued*, in contrast to the usual conventions made in real calculus.

The n values of $\sqrt[n]{z}$ can easily be determined as follows. We set

$$w = R(\cos \phi + i \sin \phi) \quad \text{and} \quad z = r(\cos \theta + i \sin \theta).$$

Then from de Moivre's formula (cf. Sec. 0.5) it follows that

$$z = w^n = R^n(\cos n\phi + i \sin n\phi) = r(\cos \theta + i \sin \theta).$$

By equating the absolute values on both sides we have

$$R^n = r \quad \text{or} \quad R = \sqrt[n]{r}$$

where the root is real positive and thus uniquely determined. By equating the arguments we obtain

$$n\phi = \theta + 2k\pi \quad \text{or} \quad \phi = \frac{\theta}{n} + \frac{2k\pi}{n}$$

where k is an integer. Consequently, $\sqrt[n]{z}$, for $z \neq 0$, has the n distinct values

$$(5) \quad \sqrt[n]{z} = \sqrt[n]{r} \left(\cos \frac{\theta + 2k\pi}{n} + i \sin \frac{\theta + 2k\pi}{n} \right), \quad k = 0, 1, \dots, n - 1.$$

These n values lie on a circle of radius $\sqrt[n]{r}$ with center at the origin and constitute the vertices of a regular polygon of n sides.

The value of $\sqrt[n]{z}$ obtained by taking the principal value of $\arg z$ (cf. Sec. 10.1) and $k = 0$ in (5) is called the **principal value of the n -valued function $w = \sqrt[n]{z}$** .

Example 1. $w = \sqrt{z}$ has the values

$$z_1 = \sqrt{r} \left(\cos \frac{\theta}{2} + i \sin \frac{\theta}{2} \right)$$

and

$$z_2 = \sqrt{r} \left[\cos \left(\frac{\theta}{2} + \pi \right) + i \sin \left(\frac{\theta}{2} + \pi \right) \right] = -z_1$$

which lie symmetric with respect to the origin. For instance,

$$\sqrt{4i} = \pm 2 \left(\cos \frac{\pi}{4} + i \sin \frac{\pi}{4} \right) = \pm (\sqrt{2} + i\sqrt{2}).$$

Example 2. If z is positive real, then $w = \sqrt[3]{z}$ has the real value $\sqrt[3]{r}$ and the conjugate complex values

$$\sqrt[3]{r} \left(\cos \frac{2\pi}{3} + i \sin \frac{2\pi}{3} \right) = \sqrt[3]{r} \left(-\frac{1}{2} + \frac{\sqrt{3}}{2}i \right)$$

and

$$\sqrt[3]{r} \left(\cos \frac{4\pi}{3} + i \sin \frac{4\pi}{3} \right) = \sqrt[3]{r} \left(-\frac{1}{2} - \frac{\sqrt{3}}{2}i \right).$$

For instance, $\sqrt[3]{1} = 1, -\frac{1}{2} \pm \frac{1}{2}\sqrt{3}i$ (Fig. 281). Clearly these are the roots of the equation $w^3 - 1 = 0$.

Example 3. From (5) it follows that the n th root of unity is given by the formula

$$\sqrt[n]{1} = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}, \quad k = 0, 1, \dots, n - 1.$$

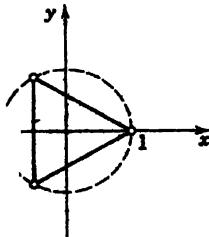


Fig. 281. $\sqrt[3]{1}$.

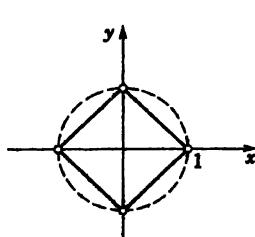


Fig. 282. $\sqrt[4]{1}$.

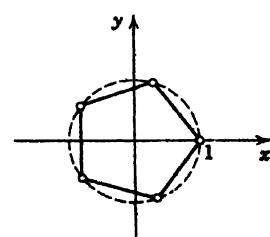


Fig. 283. $\sqrt[6]{1}$.

If ω denotes the value corresponding to $k = 1$, then the n values of $\sqrt[n]{1}$ can be written as $1, \omega, \omega^2, \dots, \omega^{n-1}$. These values are the vertices of a regular polygon of n sides inscribed in the unit circle, with one vertex at the point 1. For instance, $\sqrt[5]{1}$ has the values $1, i, -1, -i$ (Fig. 282). Fig. 283 shows $\sqrt[5]{1}$.

If w_1 is any n th root of an arbitrary complex number z , then

$$w_1, \quad w_1\omega, \quad w_1\omega^2, \quad \dots, \quad w_1\omega^{n-1}$$

are the n values of $\sqrt[n]{z}$, because the multiplication of a number w_1 by ω^k corresponds to an increase of the argument of w_1 by $2k\pi/n$.

Problems

Determine all values of the following roots and plot the corresponding points in the complex plane.

- | | | | |
|-------------------|--------------------|--------------------|-------------------------|
| 1. \sqrt{i} | 2. $\sqrt{-i}$ | 3. $\sqrt{-4}$ | 4. $\sqrt{1-i\sqrt{3}}$ |
| 5. $\sqrt[3]{-1}$ | 6. $\sqrt[3]{i}$ | 7. $\sqrt[3]{-i}$ | 8. $\sqrt[3]{1+i}$ |
| 9. $\sqrt[4]{-1}$ | 10. $\sqrt[5]{-1}$ | 11. $\sqrt[6]{-1}$ | 12. $\sqrt[8]{-1}$ |

Find and plot all solutions of the following equations.

$$13. z^3 = 27 \qquad 14. z^4 + 5z^2 = 36 \qquad 15. z^4 - (1+4i)z^2 + 4i = 0$$

16. Prove that

$$\sqrt{z} = \pm [\sqrt{|z| + x}/2 + (\text{sign } y)i\sqrt{(|z| - x)/2}] \quad (z = x + iy)$$

where $\text{sign } y = 1$ if $y \geq 0$, $\text{sign } y = -1$ if $y < 0$, and all square roots of positive numbers are taken with the positive sign. Hint: set $\sqrt{z} = w = u + iv$, square, separate into two real equations, express u^2 and v^2 in terms of x and y .

Using the result of Prob. 16, find

- | | | |
|-------------------|---------------------|--------------------|
| 17. $\sqrt{4i}$ | 18. $\sqrt{-4i}$ | 19. $\sqrt{3-4i}$ |
| 20. $\sqrt{3+4i}$ | 21. $\sqrt{-5+12i}$ | 22. $\sqrt{-8-6i}$ |

Using the result of Prob. 16, solve the following equations.

$$23. z^2 + z + 1 = i \qquad 24. z^2 - 3z + 3 = i \qquad 25. z^2 - (5+i)z + 8 + i = 0$$

10.5 Exponential Function

The real exponential function e^x has the properties

$$(1) \qquad (e^x)' = e^x$$

$$(2) \qquad e^{x_1+x_2} = e^{x_1}e^{x_2}$$

and the Maclaurin⁴ series

$$(3) \qquad e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

⁴ COLIN MACLAURIN (1698–1746), Scots mathematician.

The exponential function for complex $z = x + iy$ is denoted by e^z and is defined in terms of the real functions e^x , $\cos y$, and $\sin y$, as follows:

$$(4) \quad e^z = e^x(\cos y + i \sin y).$$

This definition is suggested by the following facts. When $z = x$ is real, then $e^z = e^x$. From the Cauchy-Riemann equations it follows that e^z is analytic for all z . From (3) in Sec. 10.3 we see that

$$(e^z)' = \frac{\partial}{\partial x} e^x \cos y + i \frac{\partial}{\partial x} e^x \sin y = e^x \cos y + ie^x \sin y,$$

that is,

$$(5) \quad (e^z)' = e^z.$$

Furthermore, setting $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$ and applying the addition formulas (7) and (8) in Sec. 0.1, we readily find

$$(6) \quad e^{z_1+z_2} = e^{z_1}e^{z_2},$$

which is the analogue of (2).

We shall see later (in Chap. 14) that complex analytic functions have Taylor series developments which are similar to the familiar Taylor series of real functions. It will then easily follow that if we replace x in (3) by z , we obtain the Maclaurin⁵ series of e^z .

From (6) with $z_1 = x$ and $z_2 = iy$ we have

$$(7) \quad e^z = e^{x+iy} = e^x e^{iy}.$$

From (4) we obtain the Euler formula

$$(8) \quad e^{iy} = \cos y + i \sin y.$$

This shows that the trigonometric representation of a complex number $z = x + iy$ may now be written

$$(9) \quad z = r(\cos \theta + i \sin \theta) = re^{i\theta}.$$

Furthermore, from (8) we see that

$$(10) \quad |e^{iy}| = \sqrt{\cos^2 y + \sin^2 y} = 1,$$

a result which the student should remember. Consequently, by (4),

$$(11) \quad |e^z| = e^x, \quad \arg e^z = y.$$

⁵ This series may be used for defining e^z . For this purpose one would first have to consider Chap. 13 on series (which is independent of Chaps. 11 and 12 and does not refer to real series). The properties of e^z follow from the series, as shown in Ex. 2, Sec. 14.2. Our present approach has the advantage that integrals (Chap. 12) will appear earlier and we shall have a closer connection between the general consideration of series (Chap. 13) and the Taylor and Laurent series (Chap. 14), for which integrals will be needed. We also avoid irritating poorer students by series at such an early stage.

Since $\cos 2\pi = 1$ and $\sin 2\pi = 0$, we obtain from (8)

$$(12) \quad e^{2\pi i} = 1.$$

Similarly,

$$(13) \quad e^{\pi i} = e^{-\pi i} = -1, \quad e^{\pi i/2} = i, \quad e^{-\pi i/2} = -i.$$

From (12) and (6) we find

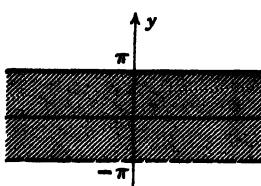


Fig. 284. Fundamental region of the exponential function e^z .

$$e^{z+2\pi i} = e^z e^{2\pi i} = e^z,$$

which shows that e^z is periodic with the imaginary period $2\pi i$. Thus

$$(14) \quad e^{z \pm 2n\pi i} = e^z \quad (n = 0, 1, \dots).$$

Because of the periodicity all the values which $w = e^z$ can assume are already assumed in the strip (Fig. 284)

$$(15) \quad -\pi < y \leq \pi.$$

This infinite strip is called a fundamental region of e^z .

The reader may prove that (Prob. 3)

$$(16) \quad e^z \neq 0 \quad \text{for all } z.$$

Problems

1. Using the Cauchy-Riemann equations, show that e^z is analytic for all z .
2. Derive (6) as explained in the text.
3. Prove (16).
4. Show that e^z and e^t are conjugates.

Find the value of e^z when z equals

$$5. \pi i/4 \quad 6. -\pi i/4 \quad 7. 3\pi i/4 \quad 8. \pi i/3 \quad 9. 1 + i \quad 10. 3 + \pi i$$

Find the real and imaginary parts of

$$11. e^{2z} \quad 12. e^{-4z} \quad 13. e^{z^2} \quad 14. e^{z^3} \quad 15. e^{z^2+z}$$

Write the following expressions in the form (9).

$$16. \sqrt{i}, \sqrt{-i} \quad 17. 3 - 4i \quad 18. \sqrt[3]{z}, \sqrt[3]{\bar{z}}$$

Find all solutions of the following equations and plot some of the solutions in the complex plane.

$$19. e^z = 1 \quad 20. e^z = 2 \quad 21. e^z = -2 \quad 22. e^{z^2} = 1$$

Show that the following functions are harmonic and find a conjugate.

$$23. u = e^{-x} \cos y \quad 24. u = e^{xy} \cos(x^2/2 - y^2/2) \quad 25. v = e^x \cos y$$

10.6 Trigonometric and Hyperbolic Functions

From Euler's formula (8) in the last section we obtain

$$\cos x = \frac{1}{2}(e^{ix} + e^{-ix}), \quad \sin x = \frac{1}{2i}(e^{ix} - e^{-ix}) \quad (x \text{ real}).$$

This suggests the following definitions for complex $z = x + iy$:

$$(1) \quad \cos z = \frac{1}{2}(e^{iz} + e^{-iz}), \quad \sin z = \frac{1}{2i}(e^{iz} - e^{-iz}).$$

Furthermore, in agreement with the familiar definitions from real calculus we define

$$(2) \quad \tan z = \frac{\sin z}{\cos z}, \quad \cot z = \frac{\cos z}{\sin z}$$

and

$$(3) \quad \sec z = \frac{1}{\cos z}, \quad \csc z = \frac{1}{\sin z}$$

Since e^z is analytic for all z , the same is true for the functions $\cos z$ and $\sin z$. The functions $\tan z$ and $\sec z$ are analytic except at the points where $\cos z$ is zero, and $\cot z$ and $\csc z$ are analytic except where $\sin z$ is zero.

The functions $\cos z$ and $\sec z$ are even, and those other functions are odd:

$$(4) \quad \begin{aligned} \cos(-z) &= \cos z & \sin(-z) &= -\sin z \\ \cot(-z) &= -\cot z & \tan(-z) &= -\tan z \end{aligned}$$

etc. Since the exponential function is periodic, the trigonometric functions are also periodic, and we have

$$(5) \quad \begin{aligned} \cos(z \pm 2n\pi) &= \cos z & \sin(z \pm 2n\pi) &= \sin z \\ \tan(z \pm n\pi) &= \tan z & \cot(z \pm n\pi) &= \cot z \end{aligned}$$

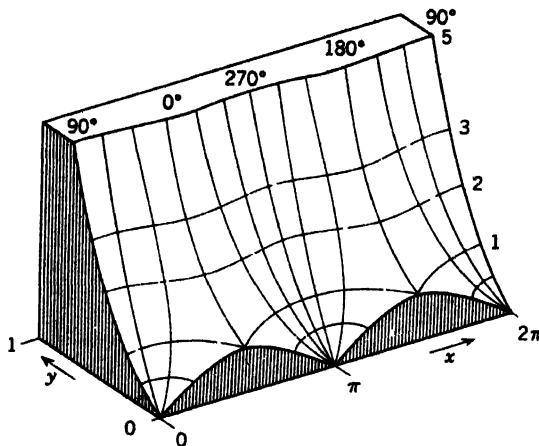
where $n = 0, 1, \dots$.

From the definitions it follows immediately that *all the familiar formulas for the real trigonometric functions continue to hold for complex values*. For example,

$$(6) \quad \frac{d}{dz} \cos z = -\sin z, \quad \frac{d}{dz} \sin z = \cos z, \quad \frac{d}{dz} \tan z = \sec^2 z.$$

$$(7) \quad \begin{aligned} \cos(z_1 \pm z_2) &= \cos z_1 \cos z_2 \mp \sin z_1 \sin z_2 \\ \sin(z_1 \pm z_2) &= \sin z_1 \cos z_2 \pm \sin z_2 \cos z_1 \\ \sin^2 z + \cos^2 z &= 1, \end{aligned}$$

and so on.

Fig. 285. Modular surface of $\sin z$.

From (1) we see that *Euler's formula is valid for complex values:*

$$(8) \quad e^{iz} = \cos z + i \sin z.$$

Using (7), we may readily represent $\cos z$ and $\sin z$ in terms of real functions. We first have

$$(9) \quad \begin{aligned} \cos(x + iy) &= \cos x \cos iy - \sin x \sin iy \\ \sin(x + iy) &= \sin x \cos iy + \cos x \sin iy. \end{aligned}$$

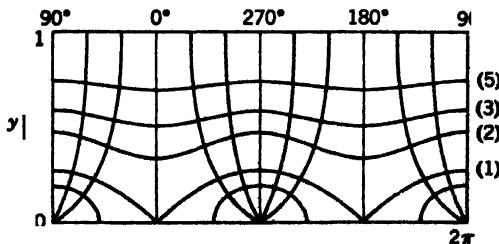
Now from (1) it follows that [cf. (16), Sec. 0.1]

$$\cos iy = \frac{1}{2}(e^{-y} + e^y) = \cosh y, \quad \sin iy = \frac{1}{2i}(e^{-y} - e^y) = i \sinh y.$$

This altogether yields the desired representations

$$(10) \quad \begin{aligned} \cos(x + iy) &= \cos x \cosh y - i \sin x \sinh y \\ \sin(x + iy) &= \sin x \cosh y + i \cos x \sinh y \end{aligned}$$

which are useful for numerical computation of $\cos z$ and $\sin z$.

Fig. 286. Curves $|\sin z| = \text{const}$ and $\arg(\sin z) = \text{const}$ in the z -plane.

The hyperbolic cosine and sine of a complex variable z are defined by the formulas

$$(11) \quad \cosh z = \frac{1}{2}(e^z + e^{-z}), \quad \sinh z = \frac{1}{2}(e^z - e^{-z}),$$

in agreement with the familiar definitions (16), Sec. 0.1, for a real variable. These functions are analytic in the entire plane.

From (11) and (1) it follows that

$$(12) \quad \cosh z = \cos(iz), \quad \sinh z = -i \sin(iz).$$

As in real calculus we define

$$(13) \quad \tanh z = \frac{\sinh z}{\cosh z}, \quad \coth z = \frac{\cosh z}{\sinh z},$$

$$(14) \quad \operatorname{sech} z = \frac{1}{\cosh z}, \quad \operatorname{csch} z = \frac{1}{\sinh z}.$$

The absolute value $|f(z)|$ of an analytic function $f(z)$, $z = x + iy$, is a real function of the two real variables x and y and can, therefore, be represented by a surface in three-dimensional space; to each point (x, y) in the xy -plane there corresponds a point with Cartesian coordinates $(x, y, |f|)$ in space, and these points form that surface which is called the **modular surface** of $f(z)$. On this surface we may draw some of the curves $|f| = \text{const}$ and $\arg f = \text{const}$; this yields a useful graphical representation of an analytic function and a geometrical illustration of its behavior.

Figure 285 shows a portion of the modular surface of $f(z) = \sin z$ and some curves $|\sin z| = \text{const}$ and $\arg(\sin z) = \text{const}$ on the surface. Figure 286 shows the orthogonal projection of these curves into the xy -plane; this graphical representation of the complex sine function is well known and used in electrical engineering.

Modular surfaces have interesting geometrical properties which are considered in Ref. [C4] in Appendix 1.

Problems

1. Prove (4). 2. Obtain (5) from (1) 3. Prove (6). 4. Prove (7).
5. Show that $\cos z$, $\sin z$, $\cosh z$, and $\sinh z$ are analytic for all z .
6. Prove (12). 7. Show that $\cos \bar{z} = \overline{\cos z}$, $\sin \bar{z} = \overline{\sin z}$.

Show that

8. $\cos z = \cosh iz$
9. $\sin z = -i \sinh iz$
10. $(\cosh z)' = \sinh z$
11. $(\sinh z)' = \cosh z$
12. $\cosh^2 z - \sinh^2 z = 1$
13. $\tanh(z + i\pi) = \tan z$
14. $\cosh z = \cosh x \cos y + i \sinh x \sin y$
15. $\sinh z = \sinh x \cos y + i \cosh x \sin y$

Find

$$\begin{array}{llll} 16. |\cos z| & 17. |\sin z| & 18. |\sec z| & 19. |\tan z| \\ 20. \operatorname{Re} \sec z & 21. \operatorname{Im} \csc z & 22. \operatorname{Re} \tan z & 23. \operatorname{Re} \cot z \end{array}$$

24. Show that all the values for which (a) $\cos z = 0$, (b) $\sin z = 0$, are real.
 25. Determine all values of z for which (a) $\cos z$, (b) $\sin z$, have real values.

Using Table 1 in Sec. 0.1, calculate

$$\begin{array}{llll} 26. e^{2+3\pi i} & 27. e^{2+i} & 28. \cos(1 + 2i) & 29. \sin(1.7 + 1.5i) \\ 30. \cos i & 31. \sin i & 32. \cosh i & 33. \sinh(2 + i) \end{array}$$

Find all solutions of the following equations.

$$\begin{array}{lll} 34. \sin z = 2 & 35. \cos z = 5 & 36. \cosh z = 0 \\ 37. \sinh z = 0 & 38. \sin z = 1000 & 39. \cosh z = 0.5 \end{array}$$

40. Verify by differentiation that $\operatorname{Re} \cos z$, $\operatorname{Im} \cos z$, $\operatorname{Re} \sin z$, and $\operatorname{Im} \sin z$ are harmonic functions.

10.7 Logarithm. General Power

The *natural logarithm* of $z = x + iy$ is denoted by $\ln z$ (sometimes also by $\log z$) and is defined as the inverse of the exponential function; that is, $w = \ln z$ is the function which satisfies the relation

$$(1) \quad e^w = z$$

for each $z \neq 0$.

Setting $w = u + iv$ and $z = |z| e^{i\theta} = re^{i\theta}$ in (1), we have

$$e^w = e^{u+i\theta} = e^u e^{iv} = re^{i\theta}.$$

It follows that

$$e^u = |z| = r \quad \text{or} \quad u = \ln |z|$$

where $\ln |z|$ is the elementary real natural logarithm of the positive number $|z|$, and

$$v = \theta = \arg z.$$

Therefore,

$$(2) \quad \ln z = \ln |z| + i \arg z = \ln \sqrt{x^2 + y^2} + i \arg(x + iy).$$

Since the argument of z is determined only up to multiples of 2π , the complex natural logarithm is infinitely many-valued.

The value of $\ln z$ corresponding to the principal value of $\arg z$, that is,

$$-\pi < \arg z \leq \pi \quad (\text{Sec. 10.1})$$

is called the **principal value** of $\ln z$ and is often denoted by $\operatorname{Ln} z$.

Obviously the other values of $\ln z$ are then of the form

$$(3) \quad \ln z = \operatorname{Ln} z \pm 2n\pi i \quad (n = 1, 2, \dots);$$

they have the same real part, and their imaginary parts differ by multiples of 2π , in agreement with the fact that e^z is periodic with the imaginary period $2\pi i$.

Also, if z is real positive, the principal value of $\arg z$ is zero, and the principal value $\ln z$ is identical with the real natural logarithm known from elementary calculus. If z is real negative, the principal value of $\arg z$ is π , and then

$$\ln z = \ln |z| + \pi i.$$

Example 1.

$$\ln(-1) = \pm\pi i, \pm 3\pi i, \pm 5\pi i, \dots, \quad \ln(-1) = \pi i$$

$$\ln i = \frac{\pi}{2} i, -\frac{3\pi}{2} i, \frac{5\pi}{2} i, -\frac{7\pi}{2} i, \frac{9\pi}{2} i, \dots, \quad \ln i = \frac{\pi}{2} i$$

$$\ln(-i) = -\frac{\pi}{2} i, \quad \ln(-2 - 2i) = \ln \sqrt{8} - \frac{3}{4}\pi i.$$

The familiar relations for the natural logarithm continue to hold for complex values, that is,

$$(4) \quad (a) \ln(z_1 z_2) = \ln z_1 + \ln z_2, \quad (b) \ln(z_1/z_2) = \ln z_1 - \ln z_2$$

but these relations are to be understood in the sense that each value of one side is also contained among the values of the other side.

Example 2. Let

$$z_1 = z_2 = e^{\pi i} = -1.$$

Taking

$$\ln z_1 = \ln z_2 = \pi i,$$

then (4a) holds provided we write $\ln(z_1 z_2) = \ln 1 = 2\pi i$; it is not true for the principal value, $\ln(z_1 z_2) = \ln 1 = 0$.

By applying (3), Sec. 10.3, to (2), this section, we find

$$\begin{aligned} \frac{d}{dz} \ln z &= \frac{\partial}{\partial x} \ln \sqrt{x^2 + y^2} + i \frac{\partial}{\partial x} (\arg z) \\ &= \frac{x}{x^2 + y^2} + i \frac{1}{1 + (y/x)^2} \left(-\frac{y}{x^2} \right) = \frac{x - iy}{x^2 + y^2} = \frac{1}{z} \end{aligned}$$

that is, the derivative of the natural logarithm is

$$(5) \quad \frac{d}{dz} \ln z = \frac{1}{z} \quad (z \neq 0).$$

Hence the principal value $\ln z$ ($z \neq 0$), which is a single-valued function, is analytic in the domain $-\pi < \arg z < \pi$ of the z -plane, that is, everywhere except at points of the negative real axis (where its imaginary part is not even continuous but has a jump of magnitude 2π).

General powers of a complex number $z = x + iy$ ($\neq 0$) are defined by the formula

$$(6) \quad z^c = e^{c \ln z} \quad (c \text{ complex}, z \neq 0).$$

Since $\ln z$ is infinitely many-valued, z^c will, in general, be multi-valued. The particular value

$$z^c = e^{c \operatorname{Ln} z}$$

is called the *principal value* of z^c .

If $c = n = 1, 2, \dots$, then z^n is single-valued and identical with the usual n th power of z . If $c = -1, -2, \dots$, the situation is similar.

If $c = 1/n$ where $n = 2, 3, \dots$, then

$$z^c = \sqrt[n]{z} = e^{(1/n) \operatorname{Ln} z} \quad (z \neq 0),$$

the exponent is determined up to multiples of $2\pi i/n$, and we obtain the n distinct values of the n th root, in agreement with the result in Sec. 10.4. If $c = p/q$, the quotient of two positive integers, the situation is similar, and z^c has only finitely many distinct values. However, if c is real irrational or genuinely complex, then z^c is infinitely many-valued.

Example 3.

$$i^n = e^{i \operatorname{Ln} i} = e^{i[(\pi/2) + 2n\pi]} = e^{-(\pi/2) + 2n\pi}$$

All these values are real, and the principal value ($n = 0$) is $e^{-\pi/2}$.

It is a *convention* that for real positive $z = x$ the expression z^c means $e^{c \operatorname{Ln} x}$ where $\operatorname{Ln} x$ is the elementary real natural logarithm (that is, the principal value of $\ln z$ ($z = x > 0$) in the sense of our definition). Also, if $z = e$, the base of the natural logarithm, $z^c = e^c$ is *conventionally* regarded as the unique value obtained from (4) in Sec. 10.5.

According to (6), the *general exponential function* with a complex constant a as its base can be written

$$(7) \quad a^z = e^{z \operatorname{Ln} a}.$$

Problems

1. Verify (4) for $z_1 = i$, $z_2 = -1$.
2. Using (7), Sec. 10.3, show that $\operatorname{Ln} z$ ($z \neq 0$) is analytic in the region $-\pi < \theta < \pi$ where θ is the principal value of $\arg z$.
3. Show that $\operatorname{Ln} z$ is not continuous on the negative real axis.
4. Show that $e^{\operatorname{Ln} z} = z$, $\operatorname{Ln}(e^z) = z \pm 2n\pi i$, $n = 0, 1, \dots$.

Determine all values of the given expressions and plot some of them in the complex plane.

$$5. \operatorname{Ln} 1 \quad 6. \operatorname{Ln} 2 \quad 7. \operatorname{Ln} i \quad 8. \operatorname{Ln} e$$

$$9. \operatorname{Ln}(ie) \quad 10. \operatorname{Ln}(-ie) \quad 11. \operatorname{Ln}(e^i) \quad 12. \operatorname{Ln}(e^{-i})$$

Solve the following equations for z .

$$13. \operatorname{Ln} z = -\pi i/2 \quad 14. \operatorname{Ln} z = \pi i/2 \quad 15. \operatorname{Ln} z = 1 + \pi i$$

Using Table 1 in Sec. 0.1, calculate the principal value $\ln z$ when z equals

16. $(1 - i)^2$

17. $\sqrt{2} + i\sqrt{2}$

18. -7

19. $3 + i\sqrt{27}$

Find the principal value of

20. $(2i)^{1/2}$

21. $(1 + i)^i$

22. $(1 + i)^{1-i}$

23. $(1 - i)^{1+i}$

The inverse sine $w = \sin^{-1} z$ is defined as the function which satisfies the relation $\sin w = z$. The inverse cosine $w = \cos^{-1} z$ is defined as the function which satisfies the relation $\cos w = z$. The other inverse trigonometric and hyperbolic functions are defined and denoted in a similar fashion. Using representations in terms of exponential functions ($\sin w = (e^{iw} - e^{-iw})/2i$, etc.), show that

24. $\sin^{-1} z = -i \ln(iz + \sqrt{1 - z^2})$

25. $\cos^{-1} z = -i \ln(z + \sqrt{z^2 - 1})$

26. $\cosh^{-1} z = \ln(z + \sqrt{z^2 - 1})$

27. $\sinh^{-1} z = \ln(z + \sqrt{z^2 + 1})$

28. $\tan^{-1} z = \frac{i}{2} \ln \frac{i+z}{i-z}$

29. $\tanh^{-1} z = \frac{1}{2} \ln \frac{1+z}{1-z}$

30. Show that $w = \sin^{-1} z$ is infinitely many-valued, and if w_1 is one of these values, the others are of the form $w_1 \pm 2n\pi$ and $\pi - w_1 \pm 2n\pi$, $n = 0, 1, \dots$. (The principal value of $w = u + iv = \sin^{-1} z$ is defined to be the value for which $-\pi/2 \leq u \leq \pi/2$ when $v \geq 0$ and $-\pi/2 < u < \pi/2$ when $v < 0$.)

Chapter 11

Conformal Mapping

If a complex function $w = f(z)$ is defined in a domain D of the z -plane, then to each point in D there corresponds a point in the w -plane. In this way we obtain a correspondence or *mapping* between D and the range of values of $f(z)$ in the w -plane. We shall see that, if $f(z)$ is an analytic function, the mapping given by $f(z)$ is *conformal* (angle-preserving), except at points where the derivative $f'(z)$ is zero.

Conformal mapping is important in engineering mathematics, because it is a standard method for solving boundary value problems in potential theory by transforming a given complicated region into a simpler one.

We shall first define and explain the concept of mapping and then consider the mappings corresponding to elementary analytic functions. Applications will be included in this chapter as well as in Chap. 16.

Prerequisites for this chapter: Chap. 10.

References: Appendix 1, Part G.

Answers to problems: Appendix 2.

11.1 *Mapping*

A continuous *real* function $y = f(x)$ of a real variable x can be exhibited graphically by plotting a curve in the Cartesian xy -plane; this curve is called the *graph* of the function. In the case of a *complex* function

$$(1) \quad w = f(z) = u(x, y) + iv(x, y) \quad (z = x + iy)$$

the situation is more complicated, because each of the complex variables w and z is represented geometrically by the points in the complex plane. This suggests the use of two separate complex planes for the two variables: one the z -plane, in which the point $z = x + iy$ is to be plotted, and the other the w -plane, in which the corresponding point $w = u + iv$ is to be plotted. In this way the function $w = f(z)$ defines a correspondence between points of these two planes. This correspondence is called a *mapping* (or *transformation*) of points in the z -plane onto points in the w -plane, and we say that $f(z)$ maps its domain of definition in the z -plane onto its range of values in the w -plane.

The point $w_0 = f(z_0)$ corresponding to a point z_0 is called the *image point* or *image* of the point z_0 with respect to the mapping defined by $f(z)$. If z moves along some curve C and $f(z)$ is continuous (not a constant), the corresponding point $w = f(z)$ will in general travel along a curve C^* in the w -plane. This curve is then called the *image* of the curve C , and the word "image" applies also to regions or other point sets.

We shall see that the properties of such mappings can be investigated by considering curves (and regions) in the z -plane and their images in the w -plane, and conversely. This will give more information about the functions than the consideration of individual points and their images.

Although two separate planes are used to represent w and z it is sometimes convenient to think of the mapping as effected in one plane and use such familiar terms as translation and rotation. For example, the mapping $w = z + 3$ may be interpreted as a translation which moves each point and configuration in the z -plane three units to the right.

To investigate the specific properties of a mapping defined by a given analytic function $w = u + iv = f(z)$, we may consider the images of the straight lines $x = \text{const}$ and $y = \text{const}$ in the w -plane. Another possibility is the study of the images of the circles $|z| = \text{const}$ and the straight lines through the origin. Conversely, we may consider the curves defined by $u(x, y) = \text{const}$ and $v(x, y) = \text{const}$ in the z -plane. These curves are called the **level curves** of u and v .

Example 1. Let us illustrate these various possibilities in the case of the mapping

$$(2) \quad w = z^2.$$

In this case the simplest procedure is to set $z = re^{i\theta}$ and $w = Re^{i\phi}$. Then

$$R = r^2 \quad \text{and} \quad \phi = 2\theta.$$

We see that circles $r = r_0 = \text{const}$ are mapped onto circles $R = r_0^2 = \text{const}$, and rays $\theta = \theta_0 = \text{const}$ onto rays $\phi = 2\theta_0 = \text{const}$. In particular, the positive real axis ($\theta = 0$) is mapped onto the positive real axis in the w -plane, and the positive imaginary axis ($\theta = \pi/2$) in the z -plane is mapped onto the negative real axis in the w -plane. The angles at the origin are doubled under the mapping. The first quadrant $0 \leq \theta \leq \pi/2$ is mapped upon the entire upper half of the w -plane (Fig. 287).

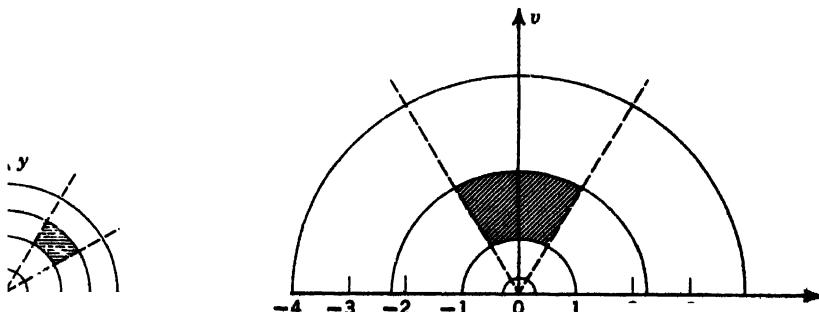


Fig. 287. Mapping $w = z^2$.

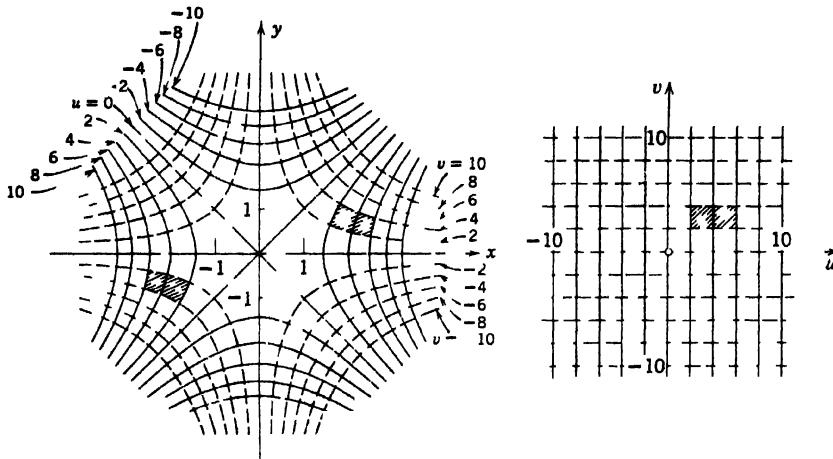


Fig. 288. Level curves of u and v in the case of the mapping $w = z^2$

In rectangular coordinates the transformation $w = z^2$ becomes

$$u + iv = x^2 - y^2 + 2xyi$$

By separating the real and the imaginary parts we obtain

$$(3) \quad u = x^2 - y^2, \quad v = 2xy$$

We see that the level curves of u and v are equilateral hyperbolas with the lines $y = \pm x$ and the coordinate axes for asymptotes. We observe that these curves are the orthogonal trajectories of each other (cf. Sec. 1.10). In Fig. 288, the two shaded domains in the z -plane are both mapped onto the shaded rectangle in the w -plane. Clearly, every point $w \neq 0$ is the image of precisely two points in the z -plane.

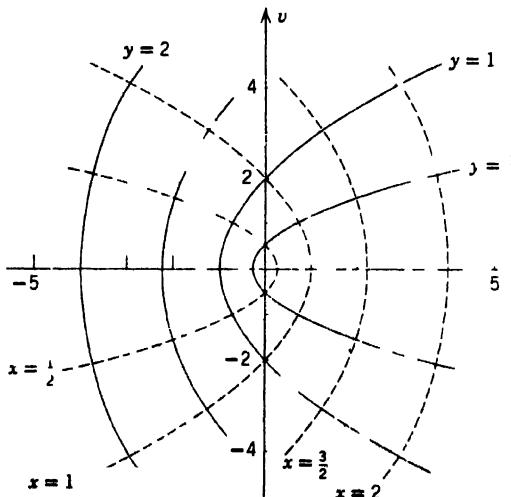


Fig. 289. Images of the lines $x = \text{const}$ and $y = \text{const}$ under the mapping $w = z^3$.

Finally, we may use (3) for determining the images of the straight lines $x = \text{const}$ and $y = \text{const}$. The line $y = c = \text{const}$ has the image

$$u = c^2 - y^2, \quad v = 2cy.$$

We may eliminate y from these equations, finding

$$v^2 = 4c^2(c^2 - u).$$

This is a parabola with focus at the origin and opening to the left. Similarly, the image of a line $y = k = \text{const}$ can be represented in the form

$$v^2 = 4k^2(k^2 + u).$$

This is a parabola with focus at the origin and opening to the right (Fig. 289).

The other powers

$$(4) \quad w = z^n, \quad n = 3, 4, \dots$$

may be considered in a similar fashion. Of course, the level curves, etc., are then represented by more complicated equations. The angular region $0 \leq \arg z \leq \pi/n$ is mapped onto the upper half of the w -plane (Fig. 290).



Fig. 290. Mapping defined by $w = z^n$.

Problems

Consider the mapping $w = u + iv = z^2$. Find and plot the images of the following curves.

- | | | |
|--------------------|---------------------|---------------------|
| 1. $y = x, y = -x$ | 2. $x = 0, 1, 2, 3$ | 3. $y = 0, 1, 2, 3$ |
| 4. $y = 1 + x$ | 5. $y = 1 - x$ | 6. $y^2 = x^2 + 1$ |

Plot the images of the following regions under the mapping $w = z^2$.

- | | | |
|--|--|-------------------------------|
| 7. $ z > 2$ | 8. $ z \leq 4$ | 9. $ \arg z < \pi/4$ |
| 10. $1 < z < 2$ | 11. $0 \leq y \leq 1$ | 12. $-\pi/4 < \arg z < \pi/2$ |
| 13. $2 < z < 3$ and $ \arg z < \pi/4$ | 14. $1/2 < z < 2$ and $\operatorname{Re} z \geq 0$ | |

15. Find and plot the images of the angular region $0 \leq \arg z \leq \pi/4$ in the case of the transformations $w = z$, $w = iz$, $w = -iz$, $w = z^2$, $w = iz^2$, $w = -z^2$, $-iz^2$, and $w = z^3$.

16. Find and plot the image of the region in Prob. 15 under the transformations $w = 1/z$, $w = i/z$, $w = 1/z^2$, and $w = i/z^2$.

17. Find and plot the image of the region $x > 0, -\pi < y < \pi$ under the transformation $w = u + iv = i(\pi^{-1}z + 1)$.

Find the images of the curves $x = \text{const}$, $y = \text{const}$, $|z| = \text{const}$, $\arg z = \text{const}$, and the level curves of u and v in the case of the following functions.

- | | | |
|---------------|---------------------|-----------------------------|
| 18. $w = iz$ | 19. $w = (2 + 2i)z$ | 20. $w = (1 + i)z/\sqrt{2}$ |
| 21. $w = 1/z$ | 22. $w = i/z$ | 23. $w = iz^2$ |

24. Find an analytic function $w = u + iv = f(z)$ which maps the half plane $z \geq 0$ onto the region $u \geq 2$ such that $z = 0$ corresponds to $w = 2 + i$.
25. Find an analytic function $w = u + iv = f(z)$ which maps the angular region $0 < \arg z < \pi/3$ onto the region $u < 1$.

11.2 Conformal Mapping

We shall now consider the most important geometrical property of the mappings defined by analytic functions, namely, their conformality.

A mapping in the plane is said to be angle-preserving, or **conformal**, if it preserves angles between oriented curves in magnitude as well as in sense, that is, the images of any two intersecting oriented curves, taken with their corresponding orientation, make the same angle of intersection as the curves, both in magnitude and direction. Here the angle between two oriented curves is defined to be the angle α ($0 \leq \alpha \leq \pi$) between their oriented tangents at the point of intersection (Fig. 291).

A curve C in the complex z -plane can be represented in the form

$$(1) \quad z = z(t) = x(t) + iy(t)$$

where t is a real parameter. For example, the function

$$z(t) = r \cos t + ir \sin t$$

represents the circle $|z| = r$, the function

$$z(t) = t + it^2$$

represents the parabola $y = x^2$, etc. The direction of increasing values of t in (1) is called the *positive direction* or *positive sense* on C . In this way (1) defines an *orientation* on C . We assume that $z(t)$ in (1) is differentiable and the derivative $\dot{z}(t)$ is continuous and nowhere zero. Then C has a unique tangent at each of its points and is called a *smooth curve*; the sense on each tangent corresponding to the positive sense on C is called the *positive sense* on that tangent, which is then oriented.

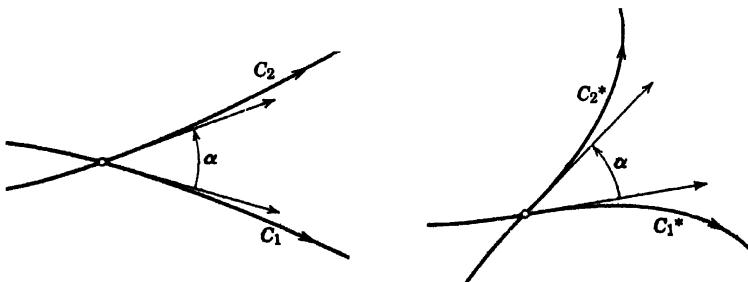


Fig. 291. Curves C_1 and C_2 and their respective images C_1^* and C_2^* under a conformal mapping.

In fact, the tangent to C at a point $z_0 = z(t_0)$ is defined as the limiting position of the straight line through z_0 and another point $z_1 = z(t_0 + \Delta t)$ as z_1 approaches z_0 along C , that is, as $\Delta t \rightarrow 0$. (Cf. also Sec. 5.13.) Now the number $z_1 - z_0$ can be represented by the vector from z_0 to z_1 (Fig. 292), and the vector corresponding to $(z_1 - z_0)/\Delta t$, where $\Delta t > 0$, has the same direction as that vector. It follows that the vector corresponding to

$$(2) \quad \dot{z}(t_0) = \frac{dz}{dt} = \lim_{\Delta t \rightarrow 0} \frac{z_1 - z_0}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{z(t_0 + \Delta t) - z(t_0)}{\Delta t}$$

is tangent to C at z_0 , and the angle between this vector and the positive x -axis is $\arg \dot{z}(t_0)$.

Consider now the mapping given by a nonconstant analytic function $w = f(z) = u(x, y) + iv(x, y)$ defined in a domain containing C . Then the image of C under this mapping is a curve C^* in the w -plane represented by

$$w = w(t) = f[z(t)].$$

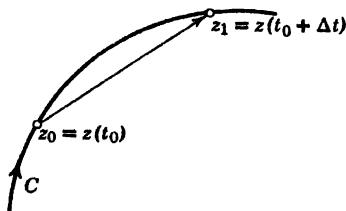


Fig. 292. Derivation of formula (2).

The point $z_0 = z(t_0)$ corresponds to the point $w(t_0)$ of C^* , and $\dot{w}(t_0)$ represents a tangent vector to C^* at this point. Now by the chain rule,

$$(3) \quad \frac{dw}{dt} = \frac{df}{dz} \frac{dz}{dt}.$$

Hence, if $f'(z_0) \neq 0$, we see that $\dot{w}(t_0) \neq 0$ and C^* has a unique tangent at $w(t_0)$, the angle between the tangent vector $\dot{w}(t_0)$ and the positive u -axis being $\arg \dot{w}(t_0)$. Since the argument of a product equals the sum of the arguments of the factors, we have from (3)

$$\arg \dot{w}(t_0) = \arg f'(z_0) + \arg \dot{z}(t_0).$$

Thus under the mapping the directed tangent to C at z_0 is rotated through the angle

$$(4) \quad \arg \dot{w}(t_0) - \arg \dot{z}(t_0) = \arg f'(z_0),$$

the angle between those two tangent vectors to C and C^* . Since the expression on the right is independent of the choice of C , we see that this angle is independent of C ; that is, the transformation $w = f(z)$ rotates the tangents of all the curves through z_0 through the same angle $\arg f'(z_0)$. Hence, two curves through z_0 which form a certain angle at z_0 are mapped upon curves forming the same angle, in sense as well as in magnitude, at the image point w_0 of z_0 . This proves the following basic result.

Theorem 1. *The mapping defined by an analytic function $f(z)$ is conformal, except at points where the derivative $f'(z)$ is zero.*

A point at which $f'(z) = 0$ is sometimes called a **critical point**. At such a point the mapping is not conformal, as may be illustrated by the following example.

Example 1. The mapping $w = z^2$ is conformal except at $z = 0$ where $w' = 2z = 0$. The conformality can be seen from the curves considered in the last section. At $z = 0$ the angles are doubled under the mapping, because each ray $\arg z = c = \text{const}$ transforms into a ray $\arg w = 2c$. Cf. Fig. 287 in the last section.

We further note that, by the definition of a derivative,

$$\lim_{z \rightarrow z_0} \left| \frac{f(z) - f(z_0)}{z - z_0} \right| = |f'(z_0)|.$$

Therefore, the mapping $w = f(z)$ magnifies the lengths of short lines by approximately the factor $|f'(z_0)|$. The image of a small figure *conforms* to the original figure in the sense that it has approximately the same shape. However, since $f'(z)$ varies from point to point, a *large* figure may have an image whose shape is quite different from that of the original figure.

We want to mention that, by (3) in Sec. 10.3 and the Cauchy-Riemann equations,

$$|f'(z)|^2 = \left| \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} \right|^2 = \left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial v}{\partial x} \right)^2 = \frac{\partial u}{\partial x} \frac{\partial v}{\partial y} - \frac{\partial u}{\partial y} \frac{\partial v}{\partial x},$$

that is,

$$(5) \quad |f'(z)|^2 = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \frac{\partial(u, v)}{\partial(x, y)}$$

where the determinant is the so-called *Jacobian* (cf. Sec. 6.3) of the transformation $w = f(z)$, written in real form,

$$u = u(x, y), \quad v = v(x, y).$$

Hence, the condition $f'(z_0) \neq 0$ implies that the Jacobian is not zero at z_0 , and it can be shown (cf. Ref. [A8] in Appendix 1) that this is a sufficient condition for the mapping given by $w = f(z)$ to be one-to-one in a sufficiently small neighborhood of z_0 . The important notion of a one-to-one mapping appearing in this connection is defined as follows.

A mapping of a region R onto a region R^* is said to be **one-to-one** if each point in R^* is the image of precisely one point in R , so that different points in R are mapped onto different points in R^* .

Example 2. The mapping $w = z^2$ is one-to-one in a sufficiently small neighborhood of any point $z \neq 0$. In a neighborhood of $z = 0$ it is not one-to-one. The full z -plane is mapped onto the w -plane so that each point $w \neq 0$ is the image of two points in the z -plane. For instance, the points $z = 1$ and $z = -1$ are both mapped onto $w = 1$, and, more generally, z_1 and $-z_1$ have the same image point $w = z_1^2$.

The practical importance of conformal mapping results from the fact that harmonic functions of two real variables (cf. Sec. 10.3) remain harmonic under a change of variables arising from a conformal transformation (Theorem 2, below). This has important consequences. Suppose that it is required to solve a **boundary value problem** in connection with a two-dimensional potential, that is, to find a solution of Laplace's equation (of two independent variables) in a given region D , assuming given values on the boundary of D . It may be possible to find a conformal mapping which transforms D into some simpler region D^* , such as a circular disk or a half-plane. Then we may solve Laplace's equation subject to the transformed boundary conditions in D^* . The resulting solution when carried back to D by the inverse transformation will be the solution of the original problem. This powerful method is justified by the following theorem.

Theorem 2. *A harmonic function $h(x, y)$ remains harmonic under a change of the variables arising from a one-to-one conformal transformation $w = f(z)$.*

Proof. Let $h(x, y)$ be harmonic in a region D . Let $w = u + iv = f(z)$ be an analytic function which maps D one-to-one and conformally onto a region D^* in the w -plane. Since the mapping is one-to-one and conformal, $f'(z) \neq 0$ in D , and the inverse function $z = f^{-1}(w)$ which maps D^* onto D exists. This function is analytic, that is, it has a derivative in D^* , namely,

$$\frac{dF}{dw} = \frac{1}{df/dz}.$$

The proof of this formula is similar as in real calculus. Let $g(x, y)$ be a conjugate harmonic function of $h(x, y)$ (cf. Sec. 10.3). Then $h + ig$ is an analytic function of $z = x + iy$, say, $H(z)$. Since the function

$$z = F(w) = x(u, v) + iy(u, v)$$

is analytic, $H[F(w)]$ is an analytic function of $w = u + iv$. Its real part $h[x(u, v), y(u, v)]$ is thus a harmonic function of u and v in D^* , and the proof is complete.

When we use the method of conformal mapping in potential theory the difficulty is to find an analytic function which maps a given region onto a simpler one. For this purpose we need some experience and a detailed knowledge of the mapping properties of the elementary analytic functions. Therefore, we shall consider the most important elementary functions from this point of view.

Problems

1. Why do the images of the curves $|z| = \text{const}$ and $\arg z = \text{const}$ under a mapping by an analytic function intersect at right angles?
2. What is the reason that the level curves $u = \text{const}$ and $v = \text{const}$ of an analytic function $w = u + iv = f(z)$ intersect at right angles at each point at which $f'(z) \neq 0$?

3. Does the mapping $w = z = x - iy$ preserve angles in size as well as in sense?

Determine the points in the z -plane at which the mapping $w = f(z)$ fails to be conformal, where

4. $w = z^4$

7. $w = e^{z^2}$

10. $w = z^2 + az + b$

5. $w = \sin z$

8. $w = \cos \pi z$

11. $w = z^4 - z^2$

6. $w = z + z^{-1} (z \neq 0)$

9. $w = z^2 + z^{-2} (z \neq 0)$

12. $w = \sin z^2$

Verify (5) for the following functions $f(z) = u(x, y) + iv(x, y)$.

13. e^z

14. $\cos z$

15. $z^2 - 3z$

16. Verify Theorem 2 for $h(x, y) = \ln(x^2 + y^2)$, $f(z) = 1/z$ ($z \neq 0$).

Represent the following curves in the z -plane ($z = x + iy$) in the form $z = z(t)$.

17. $x^2 + y^2 = 4$

18. $y = 1/x$

19. $y = 4x^2$

20. $y = 3x - 4$

11.3 Linear Transformations

We shall now consider an important class of conformal mappings, starting with the simplest types of transformations in this class.

The mapping

$$(1) \quad w = z + b$$

may be regarded as a *translation* in the direction defined by the argument of b through a distance equal to $|b|$. If $b = 0$, then we obtain the **identity transformation**

$$w = z.$$

A transformation of the form

$$(2) \quad w = az \quad (a \neq 0)$$

is a *rotation* through a fixed angle equal to $\arg a$ combined with a *uniform dilatation or contraction* such that $|w| = |a| |z|$. If a is real and positive, then (2) is a pure dilatation or contraction. If $|a| = 1$, then (2) is a pure rotation.

In the case of the transformation

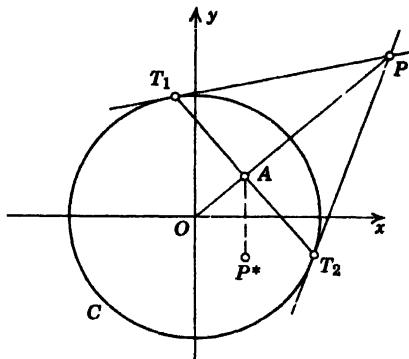
$$(3) \quad w = \frac{1}{z}$$

it is useful to introduce polar coordinates by setting

$$z = re^{i\theta} \quad \text{and} \quad w = Re^{i\phi}.$$

Then our mapping may be written

$$(3') \quad R = \frac{1}{r}, \quad \phi = -\theta.$$

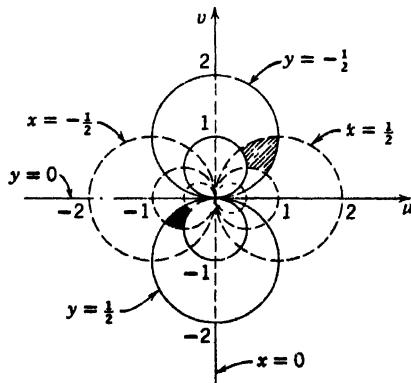
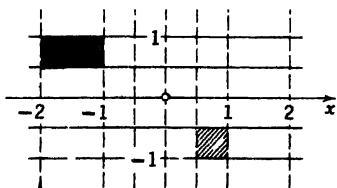
Fig. 293. Geometrical construction of $1/z$.

From this we see that a point $w = 1/z$ ($z \neq 0$) lies on the ray from the origin through z , at the distance $1/|z|$ from the origin.

We mention that the image $w = 1/z$ of a given point z can be obtained by a simple geometric construction shown in Fig. 293. In this figure, P represents the given number z . The points T_1 and T_2 are those points of the unit circle C for which the tangents to C pass through P , and A is the point of intersection of the segments OP and T_1T_2 . It can be shown by elementary methods that if OP has the length $|z|$, then OA has the length $1/|z|$. We say that A is obtained from P by **inversion with respect to the unit circle** (or "reflection" in the unit circle). Obviously, A represents the complex number $1/\bar{z}$. Hence, the point P^* , representing $w = 1/z$, is obtained by reflecting A in the real axis. This means that ***the transformation $w = 1/z$ is an inversion with respect to the unit circle combined with a reflection in the real axis.***

A fixed point of a mapping $w = f(z)$ is a point whose image is the same complex number; that is, the fixed points are obtained from

$$w = f(z) = z.$$

Fig. 294. Mapping $w = 1/z$.

The mapping $w = 1/z$ has the fixed points $z = \pm 1$, because these points are the solutions of the equation $z = 1/z$.

A transformation of the form

$$(4) \quad w = \frac{az + b}{cz + d} \quad (ad - bc \neq 0)$$

is called a *linear fractional transformation* or, briefly, a *linear transformation* or *linear mapping*. The constants a , b , c , and d in (4) are real or complex numbers.

The derivative is

$$w' = \frac{a(cz + d) - c(az + b)}{(cz + d)^2} = \frac{ad - bc}{(cz + d)^2},$$

and this shows that the condition $ad - bc \neq 0$ implies $w' \neq 0$, which entails conformality, whereas $ad - bc = 0$ leads to the uninteresting case $w' \equiv 0$ or $w = \text{const}$, which will be excluded in our further consideration.

From (4) we see that to each z for which $cz + d \neq 0$ there corresponds precisely one complex number w . Suppose that $c \neq 0$. Then to the value $z = -d/c$, for which $cz + d = 0$, there does not correspond a number w . This situation suggests that we attach an “improper point” to the w -plane. This point is called the **point at infinity** and is denoted by the symbol ∞ (*infinity*). The complex plane together with the point ∞ is called the **extended complex plane**. The complex plane without that improper point is called the **finite complex plane**. We now let $w = \infty$ be the image point of $z = -d/c$ ($c \neq 0$) under the mapping defined by (4). When $c = 0$ in (4), then $a \neq 0$ and $d \neq 0$ (why?), and we let $w = \infty$ be the image point of $z = \infty$, the improper point of the extended z -plane.

The inverse mapping of (4) is obtained by solving (4) for z ; we find

$$(5) \quad z = \frac{-dw + b}{cw - a}.$$

When $c \neq 0$, then $cw - a = 0$ for $w = a/c$, and we let $w = a/c$ correspond to $z = \infty$. When $c = 0$, then $a \neq 0$ and $d \neq 0$, and we let $w = \infty$ correspond to $z = \infty$, as before. It follows that every mapping (4) is a one-to-one mapping of the extended z -plane onto the extended w -plane; we say, that *every linear transformation (4) maps “the extended plane in a one-to-one manner onto itself.”*

The fixed points of the mapping (4) are obtained from the equation

or

$$(6) \quad cz^2 - (a - d)z - b = 0.$$

This is a quadratic equation in z whose coefficients all vanish if, and only if, the mapping is the identity (in this case, $a = d \neq 0, b = c = 0$). Hence we obtain the following result.

Theorem 1. *A linear transformation, not the identity, has at most two fixed points. If a linear transformation is known to have three or more fixed points, it must be the identity.*

Linear transformations have the following interesting property.

Theorem 2. *Any linear transformation (4) maps the totality of circles and straight lines in the z -plane onto the totality of circles and straight lines in the w -plane.*

Proof. Let $w = u + iv$. Any straight line or circle in the w -plane can be represented in the form

$$A(u^2 + v^2) + B_1u + B_2v + C = 0$$

where A, B_1, B_2 , and C are real constants. Since

$$u^2 + v^2 = |w|^2 = w\bar{w}, \quad u = \frac{1}{2}(w + \bar{w}), \quad v = \frac{1}{2i}(w - \bar{w})$$

we may write this representation in the form

$$(7) \quad Auu\bar{w} + B\dot{u}\bar{w} + \bar{B}\bar{w} + C = 0$$

where $B = \frac{1}{2}(B_1 - iB_2)$. If we insert w and \bar{w} according to (4) and multiply the resulting equation by $(cz + d)(\bar{cz} + \bar{d})$, we obtain

$$\begin{aligned} A(az + b)(az + \bar{b}) + B(az + b)(\bar{cz} + \bar{d}) \\ + \bar{B}(az + \bar{b})(cz + d) + C(cz + d)(\bar{cz} + \bar{d}) = 0. \end{aligned}$$

If we perform the multiplication, this equation becomes

$$(8) \quad \alpha z\bar{z} + \beta z + \bar{\beta}\bar{z} + \gamma = 0$$

where the coefficients are given by the formulas

$$\alpha = A|a|^2 + 2 \operatorname{Re}(Ba\bar{c}) + C|c|^2$$

$$\beta = Aab + Bad + \bar{B}bc + Ccd$$

$$\gamma = A|b|^2 + 2 \operatorname{Re}(Bbd) + C|d|^2.$$

Since α and γ are real, (8) represents a circle or a straight line in the z -plane. Conversely, starting from (8) and using the inverse transformation (5), we obtain (7), and the proof is complete.

The calculations in the proof show that a straight line may very well be mapped onto a circle, but never onto any other curve. For example, $w = 1/z$ maps a straight line or a circle through the origin onto a straight line in the w -plane, while a straight line or circle not passing through $z = 0$ is mapped onto a circle in the w -plane.

Obviously, the mappings (1), (2), and (3) are special cases of (4). Conversely, any mapping (4) is a composite of mappings of the forms (1), (2), and (3). In fact, if $c = 0$, this is immediately clear. If $c \neq 0$, then we can write (4) in the form

$$(9)^* \quad w = K \frac{1}{cz + d} + \frac{a}{c} \quad \text{where} \quad K = -\frac{ad - bc}{c}.$$

Setting

$$w_1 = cz, \quad w_2 = w_1 + d, \quad w_3 = \frac{1}{w_2}, \quad w_4 = Kw_3,$$

we have $w = w_4 + a/c$, and the statement is proved.

Special linear transformations of practical importance and further general properties of linear transformations will be considered in the following section..

Problems

1. Find $1/(4 - 3i)$ from $4 - 3i$ by means of the geometrical construction explained in connection with the transformation $w = 1/z$.

Find the fixed points of the following mappings.

2. $w = iz$	3. $w = -iz + 4$	4. $w = z^2$
5. $w = (z - i)^2$	6. $w = z^3$	7. $w = iz^2$
8. $w = \frac{3z - 1}{z + 3}$	9. $w = \frac{5z + 4}{z + 5}$	10. $w = \frac{2iz - 1}{z + 2i}$

Find a linear mapping whose fixed points are

11. $1, -1$ 12. $i, -i$ 13. 1 14. $2i$

15. Find all linear mappings whose fixed points are $-i$ and i .

Find the images of the following circles and straight lines under the mapping $w = 1/z$.

16. $ z = 1$	17. $ z + 1 = 1$	18. $ z - 1 = 1$
19. $ z - 2i = 2$	20. $y = x - 1$	21. $x = 1$

22. Prove that $w = 1/z$ maps circles onto circles (straight lines regarded as degenerate cases of circles).

23. Find the image of the region $1 < x < 2$ under the mapping $w = 1/z$.

24. Show that a mapping (4) with $c = 0$ can be represented as a composite of simpler mappings of the types (1), (2), and (3).

25. Represent $w = \frac{z + i}{iz + 4}$ as a composite of mappings of the types (1), (2), and (3).

11.4 Special Linear Transformations

In applying conformal mapping to practical problems it is often required to map the interior of the unit circle or a half-plane onto itself, or the interior

of the unit circle onto a half-plane. We want to show that this can be done by linear transformations, and these transformations can be found in a systematic way.

We start with the following *general remark* about linear transformations

$$(1) \quad w = \frac{az + b}{cz + d} \quad (ad - bc \neq 0).$$

If $z = \infty$, the right-hand side of (1) becomes $(a \cdot \infty + b)/(c \cdot \infty + d)$, which is a meaningless expression. We assign to it the value $w = a/c$ when $c \neq 0$, and $w = \infty$ when $c = 0$. This follows from our consideration of (4) and (5) in the last section.

The linear transformation (1) depends on three essential constants, namely, the ratios of any three of the constants a, b, c, d to the fourth. The requirement that three distinct points in the z -plane have specified images in the w -plane leads to a unique linear transformation, as follows.

Theorem 1. *Three given distinct points z_1, z_2, z_3 can always be mapped onto three prescribed distinct points w_1, w_2, w_3 by one, and only one, linear mapping $w = f(z)$. This mapping is given implicitly by the equation*

$$(2) \quad \frac{w - w_1}{w - w_3} \cdot \frac{w_2 - w_3}{w_2 - w_1} = \frac{z - z_1}{z - z_3} \cdot \frac{z_2 - z_3}{z_2 - z_1}.$$

(If one of these points is the point ∞ , the quotient of those two differences which contain this point must be replaced by 1.)

Proof. Equation (2) is of the form $F(w) = G(z)$ where F and G denote fractional linear functions of the respective variables. From this we obtain $w = f(z) = F^{-1}[G(z)]$ where F^{-1} denotes the inverse function of F . Since the inverse of a linear transformation and the composite of linear transformations are linear transformations (cf. Prob. 5 at the end of the section), $w = f(z)$ is a linear transformation. Furthermore, from (2) we see that

$$\begin{aligned} F(w_1) &= 0, & F(w_2) &= 1, & F(w_3) &= \infty, \\ G(z_1) &= 0, & G(z_2) &= 1, & G(z_3) &= \infty. \end{aligned}$$

Hence, $w_1 = f(z_1)$, $w_2 = f(z_2)$, $w_3 = f(z_3)$. This proves the existence of a linear transformation $w = f(z)$ which maps z_1, z_2, z_3 onto w_1, w_2, w_3 , respectively.

We prove that $w = f(z)$ is uniquely determined. Suppose that $w = g(z)$ is another linear mapping which maps z_1, z_2, z_3 onto w_1, w_2, w_3 , respectively. Then its inverse $g^{-1}(w)$ maps w_1 onto z_1 , w_2 onto z_2 , and w_3 onto z_3 . Consequently, the composite mapping $H = g^{-1}[f(z)]$ maps each of the points z_1, z_2 , and z_3 onto itself; that is, it has three distinct fixed points z_1, z_2, z_3 . From Theorem 1 in the last section it follows that H is the identity mapping, and, therefore, $g(z) = f(z)$.

The last statement of the theorem follows from the general remark at the beginning of this section. This completes the proof.

We shall now consider linear mappings of the upper half-plane $y \geq 0$ onto the unit disk $|w| \leq 1$. Clearly the x -axis, which is the boundary of that half-plane, must correspond to the unit circle $|w| = 1$ of that disk, and we have to make sure that the upper half-plane is mapped onto the interior but not onto the exterior of that unit circle.

Example 1. Find the linear transformation (1) which maps $z_1 = -1$, $z_2 = 0$, $z_3 = 1$ onto $w_1 = -1$, $w_2 = -i$, $w_3 = 1$, respectively. From (2) we obtain

$$\frac{w - (-1)}{w - 1} \cdot \frac{-i - 1}{-i - (-1)} = \frac{z - (-1)}{z - 1} \cdot \frac{0 - 1}{0 - (-1)}$$

or

$$(3) \quad w = \frac{z - i}{-iz + 1}.$$

Let us show that we can determine the specific properties of such a mapping without difficult calculations. The images of the lines $x = \text{const}$ and $y = \text{const}$ are obtained as follows. The point $z = i$ corresponds to $w = 0$, and $z = \infty$ corresponds to $w = i$. If $z = iy$, then $w = i(y - 1)/(y + 1)$; that is, the positive imaginary axis is mapped onto the segment $u = 0$, $-1 \leq v \leq 1$. Since the mapping is conformal and straight lines are mapped onto circles or straight lines, the lines $y = \text{const}$ are mapped onto circles through the image of $z = \infty$, that is, onto circles through $w = i$ and with center on the v -axis. It follows that, for the same reasons, the lines $x = \text{const}$ are mapped onto circles which are orthogonal to the image circles of the lines $y = \text{const}$ (Fig. 295). The lower half-plane corresponds to the exterior of the unit circle $|w| = 1$.

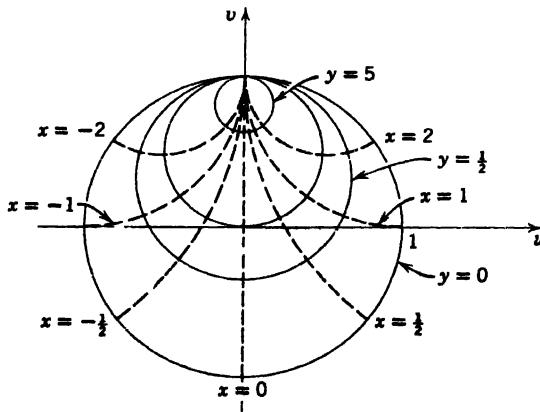


Fig. 295. Linear mapping in Example 1.

Example 2. Determine the linear transformation which maps $z_1 = 0$, $z_2 = 1$, $z_3 = \infty$ onto $w_1 = i$, $w_2 = -1$, $w_3 = -i$, respectively.

From (2) we find that the desired mapping is

$$(4) \quad \frac{z - i}{z + i},$$

in this case we have to replace the quotient $(1 - \infty)/(z - \infty)$, which first occurs, by 1.

We shall now consider *linear mappings of the upper half-plane* $y \geq 0$ onto the upper half-plane $v \geq 0$. In this case the x -axis must be mapped onto the u -axis.

Example 3. Find the linear transformation which maps the points $z_1 = -2$, $z_2 = 0$, $z_3 = 2$ onto the points $w_1 = \infty$, $w_2 = 1/4$, $w_3 = 3/8$, respectively. From (2) we obtain

$$(5) \quad \begin{aligned} z + 1 \\ \cdots \\ 2z + 4 \end{aligned}$$

We leave the discussion of the properties of this mapping to the reader (Prob. 3).

As another important class of linear mappings let us consider the *transformations which map the unit disk in the z -plane onto the disk in the w -plane*. We may readily verify that the function

$$(6) \quad w = \frac{z - z_0}{cz - 1}, \quad c = \bar{z}_0, \quad |z_0| < 1$$

belongs to this class and maps the point z_0 onto $w = 0$ (cf. Prob. 4).

Example 4. Let $z_0 = 1/2$. Then by (6),

$$w(z) = \frac{2z - 1}{z - 2}.$$

The real axes correspond to each other; in particular,

$$w(-1) = 1, \quad w(0) = 1/2, \quad w(1) = -1.$$

Since the mapping is conformal and straight lines are mapped onto circles or straight lines and $w(\infty) = 2$, the images of the lines $x = \text{const}$ are circles through $w = 2$ with centers on the u -axis; the lines $y = \text{const}$ are mapped onto circles which are orthogonal to the aforementioned circles (Fig. 296).

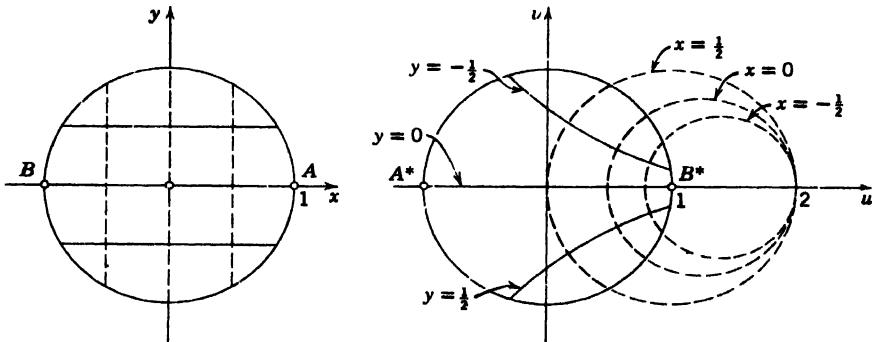


Fig. 296. Mapping in Example 4.

By combining linear transformations and transformations of the form $w = z^n$ ($n > 1$ and integral) we may obtain *mappings of angular regions onto the unit disk*.

Example 5. Map the angular region $D: -\pi/6 \leq \arg z \leq \pi/6$ onto the unit disk $|w| \leq 1$. We may proceed as follows. The mapping $t = z^3$ maps D onto the right half of the t -plane.

Then we may apply a linear transformation which maps this half-plane onto the unit disk, for example, the transformation

$$w = i \frac{t - 1}{t + 1}.$$

By inserting $t = z^3$ into this mapping we find

$$w = i \frac{z^3 - 1}{z^3 + 1};$$

this mapping has the required properties (Fig. 297).

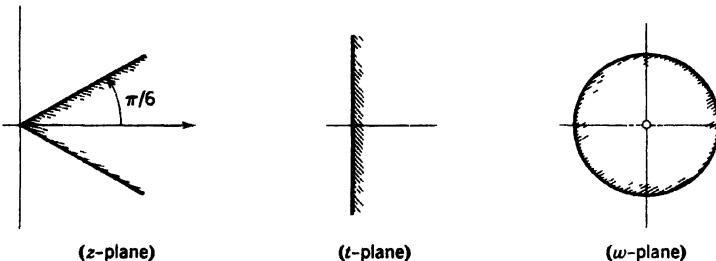


Fig. 297. Mapping in Example 5.

Problems

- Derive (4) from (2). Graph the curves corresponding to the lines $u = \text{const}$ and $v = \text{const}$.
- Find the inverse of (4) and graph the images of the lines $x = \text{const}$ and $y = \text{const}$.
- Derive (5) from (2). Find the inverse of (5) and graph the curves corresponding to $u = \text{const}$ and $v = \text{const}$.
- Prove the statement involving (6).
- Show that the composite of two linear transformations is a linear transformation.
- Find the inverse of (3). Show that (3) maps the lines $x = c = \text{const}$ onto circles with centers on the line $v = 1$.

Find the linear mapping $w = w(z)$ which maps

7. $1, i, -i$ onto $1, -i, i$, respectively.
8. $-2, 0, 2$ onto $-1/2, \infty, 1/2$, respectively.
9. $0, 1, \infty$ onto $\infty, 1, 0$, respectively.
10. $0, 1, i$ onto $2, 3, 2 + i$, respectively.
11. $0, 1, 2$ onto $1, 1/2, 1/3$, respectively.
12. $-1, 0, 1$ onto $0, -1, \infty$, respectively.
13. $-i, 0, i$ onto $\infty, -1, 0$, respectively.
14. $0, i, 2i$ onto $\infty, 2i, 5i/2$, respectively.
15. $-1, i, 1 + i$ onto $0, \infty, 2 + i$, respectively.
16. $i, 0, 1$ onto $i/2, 0, (1+i)/2$, respectively.
17. $1/2, 1, 3$ onto $\infty, 4, 6/5$, respectively.
18. $0, 1, \infty$ onto $\infty, (1-i)/2, 1/2$, respectively.

19. 1, 2, 3 onto 5, $7/4$, $9/7$, respectively.

20. Find an analytic function $w = f(z)$ which maps the region $2 \leq y \leq x + 1$ onto the unit disk $|w| \leq 1$.

11.5 Mapping by Other Elementary Functions

We shall now consider the mapping properties of some further important special functions.

The exponential function (Sec. 10.5)

$$(1) \quad w = e^z$$

defines a mapping which is conformal everywhere, because its derivative is different from zero at any point. If we set $w = Re^{i\phi}$, then

$$Re^{i\phi} = e^{x+i\phi} = e^x e^{iy},$$

and (1) can be written in the form

$$(2) \quad R = e^x, \quad \phi = y.$$

From this we see that the lines $x = a = \text{const}$ are mapped onto the circles $R = e^a$, and the lines $y = c$ are mapped onto the rays $\phi = c$. Since $e^z \neq 0$ for all z , the point $w = 0$ is not an image of any point z . A rectangular region, say, $a \leq x \leq b$, $c \leq y \leq d$, is mapped onto the region

$$e^a \leq R \leq e^b, \quad c \leq \phi \leq d$$

bounded by portions of rays and circles (Fig. 298).

The fundamental strip $-\pi < y \leq \pi$ is mapped upon the full w -plane (cut along the negative real axis). More generally, every horizontal strip bounded by two lines $y = c$ and $y = c + 2\pi$ is mapped upon the full w -plane. This illustrates the fact that e^z is periodic with the period $2\pi i$.

The horizontal strip $0 \leq y \leq \pi$ is mapped onto the upper half of the w -plane. The boundary $y = 0$ is mapped onto the positive half of the u -axis, and the line $y = \pi$ onto the negative half of the u -axis, as follows from (2). The segment from 0 to πi is mapped onto the semicircle $|w| = 1$, $v \geq 0$. The

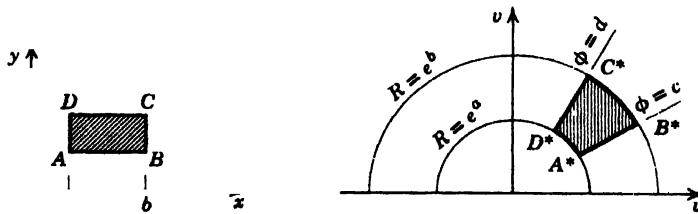
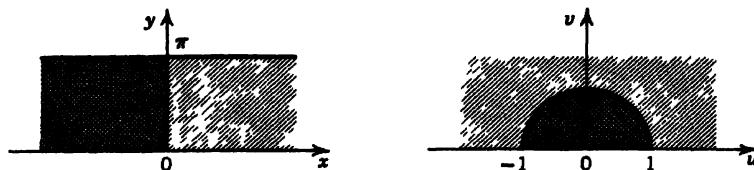


Fig. 298. Mapping by $w = e^z$.

Fig. 299. Mapping by $w = e^z$.

left half ($x \leq 0$) of our strip is mapped onto the region $|w| \leq 1, v \geq 0$, and the right half ($x \geq 0$) of the strip is mapped onto the exterior of that semi-circle $|w| = 1$ in the upper half of the w -plane (Fig. 299).

Since the logarithmic function $w = u + iv = \ln z$ is the inverse of the exponential function, the properties of the corresponding conformal mapping can be easily obtained from those of the exponential function by interchanging the roles of the z and the w planes in the preceding considerations. The principal value $w = \ln z$ thus maps the z -plane (cut along the negative real axis) onto the horizontal strip $-\pi < v \leq \pi$ of the w -plane. Further details of the mapping will be discussed in Ex. 2 of the next section.

The sine function (Sec. 10.6)

$$(3) \quad w = u + iv = \sin z = \sin x \cosh y + i \cos x \sinh y$$

where

$$(4) \quad u = \sin x \cosh y, \quad v = \cos x \sinh y,$$

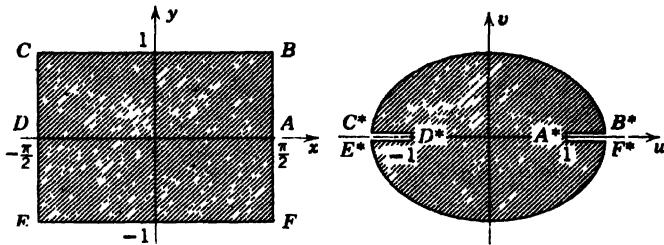
is periodic. Hence, the mapping (4) is certainly not one-to-one if we consider it in the full xy -plane. We restrict our consideration to the infinite strip D defined by $-\pi/2 \leq x \leq \pi/2$. Since $f'(z) = \cos z$ is zero at $z = \pm\pi/2$, the mapping is not conformal at these two points. From (4) we see that the boundary of D is mapped onto the u -axis. The segment $-\pi/2 \leq x \leq \pi/2$ of the x -axis maps onto the segment $1 \leq u \leq 1$ of the u -axis, the line $x = -\pi/2$ maps onto $u \leq -1, v = 0$, and the line $x = \pi/2$ maps onto $u \geq 1, v = 0$. The line segment $y = c > 0, -\pi/2 \leq x \leq \pi/2$ maps onto the portion of the ellipse

$$u = \cosh c \sin x, \quad v = \sinh c \cos x$$

or

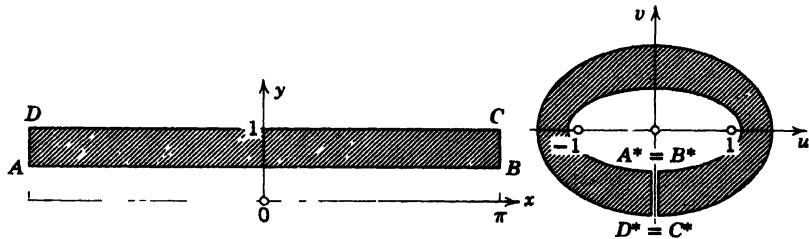
$$(5) \quad \frac{u^2}{\cosh^2 c} + \frac{v^2}{\sinh^2 c} = 1$$

in the upper half of the w -plane. The line segment $y = -c$ ($c > 0$), $-\pi/2 \leq x \leq \pi/2$ maps onto the lower half of the ellipse (5). The foci of the ellipse are at $w = \pm 1$, and we see that they are independent of c . Consequently, if we let c vary, we obtain a family of confocal ellipses. The rectangular region D defined by $-\pi/2 < x < \pi/2, -c < y < c$ is thus mapped onto the interior of the ellipse (5); but note that the image of the boundary consists of the ellipse and the two segments of the x -axis, as shown

Fig. 300. Mapping by $w = \sin z$.

in Fig. 300 (where $c = 1$). The image points of points on the vertical parts of the boundary coincide in pairs. In particular, $B^* = F^*$ and $C^* = E^*$.

Similarly, the rectangle $-\pi < x < \pi$, $c < y < d$ maps onto an elliptic ring cut along the negative v -axis (Fig. 301). The straight lines $x = \text{const}$ ($-\pi/2 < x < \pi/2$) map onto confocal hyperbolas which intersect those ellipses at right angles, and the y -axis maps onto the v -axis.

Fig. 301. Mapping by $w = \sin z$.

The cosine function

$$(6) \quad w = \cos z = \sin \left(z + \frac{\pi}{2} \right)$$

defines the same mapping as $\sin z$, preceded by a translation to the right through $\pi/2$ units.

The hyperbolic function

$$(7) \quad w = \sinh z = -i \sin(iz)$$

defines a transformation which is a rotation $t = iz$ followed by the mapping $p = \sin t$ and another rotation $w = -ip$.

Similarly, the transformation

$$(8) \quad w = \cosh z = \cos(iz)$$

is a rotation $t = iz$ followed by the mapping $w = \cos t$.

Example 1. Find the image of the semi-infinite strip $x \geq 0$, $0 \leq y \leq \pi$ (Fig. 302) under the mapping (8). We set $w = u + iv$. Since $\cosh 0 = 1$, the point $z = 0$ is mapped onto $w = 1$. For real $z = x \geq 0$, $\cosh z$ is real and increases monotone from 1 as x increases.

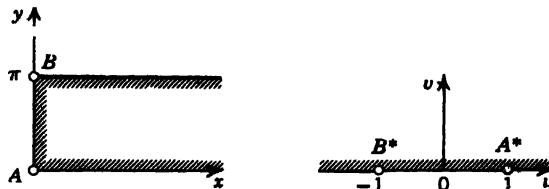


Fig. 302. Mapping in Example 1.

Hence the positive x -axis is mapped onto the portion $u \geq 1$ of the u -axis. For purely imaginary $z = iy$ we have $\cosh iy = \cos y$. It follows that the left boundary of the strip is mapped onto the segment $1 \geq u \geq -1$ of the u -axis, the point $z = \pi i$ corresponding to $w = \cosh i\pi = \cos \pi = -1$.

On the upper boundary of the strip, $y = \pi$, and since $\sin \pi = 0$, $\cos \pi = -1$, it follows that this part of the boundary is mapped onto the portion $u \leq -1$ of the u -axis. Hence the boundary of the strip is mapped onto the u -axis. It is not difficult to see that the interior of the strip is mapped onto the upper half of the w -plane, and the mapping is one-to-one.

Example 2. Find the temperature $T(x, y)$ in the strip considered in Ex. 1, if the temperature on the boundary is

$$T = T_0 \text{ on the segment from } 0 \text{ to } \pi i,$$

$$T = 0 \text{ on the upper and lower boundaries.}$$

Since T does not depend on time ("steady state temperature distribution"), the heat equation reduces to Laplace's equation (cf. Sec. 9.11)

$$\nabla^2 T = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0,$$

and we have to find a solution of this equation satisfying those boundary conditions.

To solve this problem we map the strip by means of (8) onto the upper half of the w -plane. Since the segment $0 \leq y \leq \pi i$ of the y -axis is mapped onto the segment $-1 \leq u \leq 1$ of the u -axis, the boundary conditions in the w -plane are (Fig. 303):

$$T = T_0 \text{ on the segment from } -1 \text{ to } 1,$$

$$T = 0 \text{ on the other parts of the } u \text{ axis}$$

The real and imaginary parts of analytic functions are solutions of Laplace's equation, and we have to find such a solution $T(u, v)$ in the upper half of the w -plane which satisfies those boundary conditions. For this purpose we consider the functions

$$(9*) \quad \ln(w+1) = \ln|w+1| + i\phi_1, \quad \phi_1 = \arg(w+1) = \arctan \frac{v}{u+1}$$

$$\ln(w-1) = \ln|w-1| + i\phi_2, \quad \phi_2 = \arg(w-1) = \arctan \frac{v}{u-1}$$

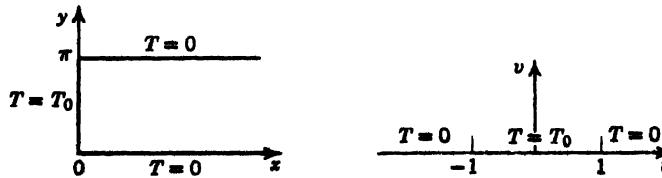


Fig. 303. Boundary conditions in Example 2.

Since $\phi_1(u, v)$ and $\phi_2(u, v)$ are harmonic functions, $\phi_2 - \phi_1$ is harmonic. If $w = u$ is real and smaller than -1 , then $\phi_2 - \phi_1 = \pi - \pi = 0$; for real $w = u$ in the interval $-1 < u < 1$ we have $\phi_2 - \phi_1 = \pi - 0 = \pi$, and for real $w = u > 1$ we have $\phi_2 - \phi_1 = 0 - 0 = 0$. Hence the function

$$(9) \quad T(u, v) = \frac{T_0}{\pi} (\phi_2 - \phi_1)$$

is harmonic in the half-plane $v > 0$ and satisfies those boundary conditions in the w -plane. Since $\tan \phi_1 = v/(u + 1)$ and $\tan \phi_2 = v/(u - 1)$, it follows that

$$\tan(\phi_2 - \phi_1) = \frac{\tan \phi_2 - \tan \phi_1}{1 + \tan \phi_1 \tan \phi_2} = \frac{2v}{u^2 + v^2 - 1},$$

and (9) takes the form

$$(10) \quad T(u, v) = \frac{T_0}{\pi} \operatorname{arc tan} \frac{2v}{u^2 + v^2 - 1}.$$

The function $w = \cosh z$ maps the strip under consideration onto the half-plane $v \geq 0$, and we have

$$w = u + iv = \cosh(x + iy) = \cosh x \cos y + i \sinh x \sin y.$$

Separating the real and imaginary parts on both sides, we may write

$$u = \cosh x \cos y, \quad v = \sinh x \sin y.$$

From these expressions for u and v it follows that in (10),

$$u^2 + v^2 - 1 = \cosh^2 x \cos^2 y + \sinh^2 x \sin^2 y - 1 = \sinh^2 x - \sin^2 y.$$

By inserting this and the expression for v into (10) and denoting $T(u(x, y), v(x, y))$ by $T^*(x, y)$ we have

$$T^*(x, y) = \frac{T_0}{\pi} \operatorname{arc tan} \frac{2 \sinh x \sin y}{\sinh^2 x - \sin^2 y}.$$

Noting that the numerator and the denominator are the imaginary part and the real part of the function $(\sinh x + i \sin y)^2$, we may write

$$T^*(x, y) = \frac{T_0}{\pi} \arg [(\sinh x + i \sin y)^2] = \frac{2T_0}{r} \arg (\sinh x + i \sin y).$$

Hence the solution of our problem is

$$(11) \quad T^*(x, y) = \frac{2T_0}{\pi} \operatorname{arc tan} \frac{\sin y}{\sinh x}$$

This function is harmonic in the interior of our strip (cf. Theorem 2 in Sec. 11.2) and it satisfies the boundary conditions. Indeed, $T^* = 0$ when $y = 0$ or $y = \pi$, and $T^* = T_0$ when $x = 0$. The isotherms (= curves of constant temperature) are the curves

$$\frac{\sin y}{\sinh x} = \text{const.}$$

In Ex. 2 we transformed the real potential $T(u, v)$ into the real potential $T^*(x, y)$ by using the function $w = u(x, y) + iv(x, y)$ which maps the half-plane onto the given region. In many cases problems of this type become simpler by working with a **complex potential**, that is, by taking a complex analytic function $F(w)$ such that the real potential $T(u, v)$ is the real or imaginary part of $F(w)$, and transforming F instead of T . It is clear that such a

complex potential F may be readily obtained from T by determining a conjugate harmonic function of T (cf. at the end of Sec. 10.3). Let us illustrate this "method of complex potentials" in the case of our previous example. A detailed consideration of complex potentials will be presented later (in Chap. 16).

Example 3. From (9*) we see that the real potential

$$T(u, v) = \frac{T_0}{\pi}(\phi_2 - \phi_1)$$

in Ex. 2 [cf. (9)] is the imaginary part of the complex potential

$$(12) \quad F(w) = \frac{T_0}{\pi}[\ln(w-1) - \ln(w+1)] = \frac{T_0}{\pi}\ln\frac{w-1}{w+1}.$$

The mapping function in Ex. 2 is

$$w = \cosh z = \frac{1}{2}(e^z + e^{-z}).$$

From this we see that in (12),

$$\frac{w-1}{w+1} = \frac{\cosh z - 1}{\cosh z + 1} = \frac{e^z + e^{-z} - 2}{e^z + e^{-z} + 2} = \frac{(e^{z/2} - e^{-z/2})^2}{(e^{z/2} + e^{-z/2})^2} = \tanh^2 \frac{z}{2}.$$

By inserting this into (12) and denoting $F(w(z))$ by $F^*(z)$ we have

$$(13) \quad F^*(z) = \frac{T_0}{\pi}\ln\tanh^2 \frac{z}{2} = \frac{2T_0}{\pi}\ln\tanh \frac{z}{2}.$$

This is the complex potential in the strip in Ex. 2, and its imaginary part is the solution of our problem. This solution may be obtained as follows.

Denoting $\tanh \frac{z}{2}$ by $H(z)$ and using (2) in Sec. 10.7, we first have

$$(14) \quad F^*(z) = \frac{2T_0}{\pi}\ln H = \frac{2T_0}{\pi}\left(\ln|H| + i\arctan\frac{\operatorname{Im} H}{\operatorname{Re} H}\right).$$

To determine the real and imaginary part of H we may start from

$$H = \tanh \frac{z}{2} = \frac{\sinh \frac{z}{2}}{\cosh \frac{z}{2}} = \frac{\sinh \frac{z}{2} \overline{\cosh \frac{z}{2}}}{\cosh \frac{z}{2} \overline{\cosh \frac{z}{2}}}.$$

By using the definitions of the hyperbolic sine and cosine in terms of exponential functions, noting that $\overline{\cosh(z/2)} = \cosh(\bar{z}/2)$, and carrying out the multiplications we find

$$H = \frac{\sinh x + i \sin y}{\cosh x + \cos y}.$$

From this we readily obtain

$$(15) \quad \operatorname{Re} H = \frac{\sinh x}{\cosh x + \cos y}, \quad \operatorname{Im} H = \frac{\sin y}{\cosh x + \cos y}.$$

By inserting these expressions into (14) we have

$$T^*(x, y) = \operatorname{Im} F^*(z) = \frac{2T_0}{\pi} \arctan \frac{\sin y}{\sinh x},$$

in agreement with our result in Ex. 2.

Furthermore, from (15) it follows that

$$|H|^2 = \frac{\sinh^2 x + \sin^2 y}{(\cosh x + \cos y)^2}.$$

and from (14) we see that the real part of $F^*(z)$ becomes

$$S^*(x, y) = \operatorname{Re} F^*(z) = \frac{T_0}{\pi} \ln \frac{\sinh^2 x + \sin^2 y}{(\cosh x + \cos y)^2}.$$

The curves $S^* = \text{const}$ intersect the isotherms $T^* = \text{const}$ at right angles and, therefore are the curves along which the heat flows.

Problems

Find and graph the images of the following regions under the mapping $w = z^3$.

$$1. |z| \leq 2, \quad 0 \leq \arg z \leq \pi/4 \quad 2. 1/2 < |z| < 2, \quad \pi/4 \leq \arg z \leq \pi/2$$

Find and graph the images of the following regions under the mapping $w = e^z$.

$$\begin{array}{ll} 3. -1 < x < 1, \quad -\pi/2 < y < \pi/2 & 4. 0 < x < 2, \quad 0 < y < \pi \\ 5. 0 < x < 1, \quad 0 < y < 1 & 6. -3 < x < -2, \quad 0 < y < \pi/4 \\ 7. -2 \leq x \leq 2, \quad -\pi \leq y \leq -\pi/2 & 8. -1 \leq x \leq 3, \quad -\pi \leq y \leq \pi \end{array}$$

9. Find an analytic function which maps the region R bounded by the positive x and y axes and the hyperbola $xy = \pi/2$ in the first quadrant onto the upper half-plane.
Hint: first map R onto a horizontal strip.

Find and graph the images of the following regions under the mapping $w = \sin z$.

$$\begin{array}{ll} 10. 0 < x < \pi/2, \quad 0 < y < 2 & 11. -\pi/2 < x < \pi/2, \quad 1 < y < 2 \\ 12. -\pi/2 < x < \pi/2, \quad 0 < y < 1 & 13. 0 < x < 2\pi, \quad 1 < y < 2 \end{array}$$

14. Find and plot the images of the lines $x = 0, \pm\pi/6, \pm\pi/3, \pm\pi/2$ under the mapping $w = \sin z$.

15. Determine all the points where the mapping $w = \sin z$ is not conformal.

16. Describe the transformation $w = \cosh z$ in terms of the transformation $w = \sin z$ and rotations and translations.

17. Verify by differentiation that (11) is harmonic.

18. Find the temperature $T(x, y)$ in the strip in Ex. 1, assuming that $T = 0$ on the left boundary and $T = 1$ on the upper and lower boundaries of the strip.

19. Show that $w = \ln \frac{z-1}{z+1}$ maps the upper half plane onto the horizontal strip $0 \leq \operatorname{Im} w \leq \pi$ as shown in Fig. 304.

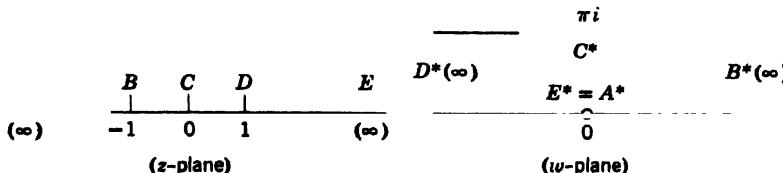


Fig. 304. Problem 19.

20. Find the image of the region $0 < \arg z < \pi/4$ under the mapping $w = \sqrt{z^2 - 1}$ (where w is the principal value of the root).

11.6 Riemann Surfaces

We consider the mapping defined by

$$(1) \quad w = u + iv = z^2$$

(cf. Sec. 11.1). This mapping is conformal, except at $z = 0$ where $w' = 2z$ is zero. The angles at $z = 0$ are doubled under the mapping. The right half of the z -plane (including the positive y -axis) maps onto the full w -plane cut along the negative half of the u -axis; the mapping is one-to-one. Similarly, the left half of the z -plane (including the negative y -axis) is mapped onto the cut w -plane in a one-to-one manner.

Obviously, the mapping of the full z -plane is not one-to-one, because every point $w \neq 0$ corresponds to precisely two points z . In fact, if z_1 is one of these points, then the other is $-z_1$. For example, $z = i$ and $z = -i$ have the same image, namely, $w = -1$, etc. Hence, the w -plane is “covered twice” by the image of the z -plane. We say that the full z -plane is mapped onto the *doubly covered* w -plane. We can still give our imagination the necessary support as follows.

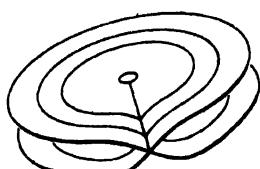


Fig. 305. Example of a Riemann surface.

We imagine one of the two previously obtained copies of the cut w -plane to be placed upon the other such that the upper sheet is the image of the right half of the z -plane, and the lower sheet is the image of the left half of the z -plane; we denote these half-planes by R and L , respectively. When we pass from R to L , the corresponding image point should pass from the upper to the lower sheet. For this reason we join the two sheets crosswise along the cut, that is, along the negative real axis. (This construction can be carried out only in imagination, since the penetration of the two sheets of a material model can only be imperfectly realized.) The two origins are fastened together. The configuration thus obtained is called a **Riemann surface**. On it every point $w \neq 0$ appears twice, at superposed positions, and the origin appears precisely once. The function $w = z^2$ now maps the full z -plane onto this Riemann surface in a one-to-one manner, and the mapping is conformal, except for the “winding point” or **branch point** at $w = 0$ (Fig. 305). This branch point is said to be of the *second order*, since it connects two sheets.

We now consider the double-valued function

$$(2) \quad w = \sqrt{z}.$$

To each $z \neq 0$ there correspond two values w , one of which is the principal value. If we replace the z -plane by the two-sheeted Riemann surface just considered, then each complex number $z \neq 0$ is represented by two points

of the surface at superposed positions. We let one of these points correspond to the principal value—for example, the point in the upper sheet—and the other to the other value. Then our function becomes a single-valued function of the points of the Riemann surface, and to any continuous motion of z on the surface there corresponds a continuous motion of the corresponding point in the w -plane. The function maps the sheet corresponding to the principal value on to the right half of the w -plane and the other sheet onto the left half of the w -plane.

Let us consider some further important examples.

Example 1. In the case of the function

$$(3) \quad w = \sqrt[n]{z} \quad n = 3, 4, \dots$$

we need a Riemann surface consisting of n sheets and having a branch point of order n at $z = 0$. One of the sheets corresponds to the principal value and the other $n - 1$ sheets to the other $n - 1$ values of the function.

Example 2. The Riemann surface of the infinitely many-valued function

$$(4) \quad w = \ln z = \operatorname{Ln} z + 2n\pi i \quad (n = 0, \pm 1, \pm 2, \dots, z \neq 0)$$

consists of infinitely many sheets. The single-valued function $w = \operatorname{Ln} z$ corresponds to one of these sheets. On this sheet the argument θ of z ranges in the interval $-\pi < \theta \leq \pi$ (cf. Sec. 10.7). The sheet is cut along the negative ray of the real axis, and the upper edge of the slit is joined to the lower edge of the next sheet which corresponds to the interval $\pi < \theta \leq 3\pi$, that is, to the single-valued function $w = \operatorname{Ln} z + 2\pi i$. In this way each value of n in (4) corresponds to precisely one of these infinitely many sheets. The function $w = \operatorname{Ln} z$ maps the corresponding sheet onto the horizontal strip $-\pi < v \leq \pi$ in the w -plane. The next sheet is mapped onto the neighboring strip $\pi < v \leq 3\pi$, etc. The function $w = \ln z$ thus maps all the sheets of the corresponding Riemann surface onto the entire w -plane, the correspondence between the points $z \neq 0$ of the Riemann surface and those of the w -plane being one-to-one.

Example 3. Let us consider the mapping defined by

$$(5) \quad w = z + \frac{1}{z} \quad (z \neq 0)$$

which is important in aerodynamics (see below). Since

$$w' = 1 - \frac{1}{z^2} = \frac{(z+1)(z-1)}{z^2}$$

the mapping is conformal except at the points $z = 1$ and $z = -1$; these points correspond to $w = 2$ and $w = -2$, respectively. From (5) we find

$$(6) \quad z = \frac{w}{2} \pm \sqrt{\frac{w^2}{4} - 1} = \frac{w}{2} \pm \frac{i}{2}\sqrt{(w+2)(w-2)}.$$

Hence, the points $w = 2$ and $w = -2$ are branch points of the second order of this function $z = z(w)$. To any value w ($\neq 2, \neq -2$) there correspond two values of z . Consequently, (5) maps the z -plane onto a two-sheeted Riemann surface, the two sheets being connected crosswise from $w = -2$ to $w = 2$ (Fig. 306), and this mapping is one-to-one. We set $z = re^{i\theta}$ and determine the images of the curves $r = \text{const}$ and $\theta = \text{const}$. From (5) we obtain

$$w = u + iv = re^{i\theta} + \frac{1}{r}e^{-i\theta} = \left(r + \frac{1}{r}\right)\cos\theta + i\left(r - \frac{1}{r}\right)\sin\theta.$$

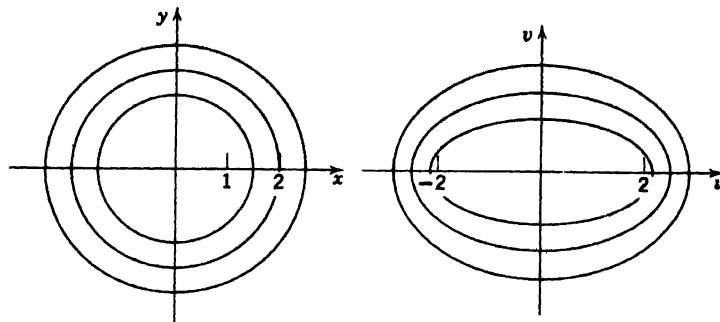


Fig. 306. Example 3.

By equating the real and imaginary parts on both sides we have

$$(7) \quad u = \left(r + \frac{1}{r} \right) \cos \theta, \quad v = \left(r - \frac{1}{r} \right) \sin \theta.$$

From this we find

$$\frac{u^2}{a^2} + \frac{v^2}{b^2} = 1, \quad \text{where} \quad a = r + \frac{1}{r}, \quad b = \boxed{\quad}$$

The circles $r = \text{const}$ are thus mapped onto ellipses whose principal axes lie in the u and v axes and have the lengths $2a$ and $2b$, respectively. Since $a^2 - b^2 = 4$, independent of r , these ellipses are confocal, with foci at $w = -2$ and $w = 2$. The unit circle $r = 1$ maps onto the line segment from $w = -2$ to $w = 2$. For every $r \neq 1$ the two circles with radii r and $1/r$ map onto the same ellipse in the w -plane, corresponding to the two sheets of the Riemann surface. Hence, the interior of the unit circle $|z| = 1$ corresponds to one sheet, and the exterior to the other.

Furthermore, from (7) we obtain

$$(8) \quad \frac{u^2}{\cos^2 \theta} - \frac{v^2}{\sin^2 \theta} = -4.$$

The lines $\theta = \text{const}$ are thus mapped onto the hyperbolas which are the orthogonal trajectories of those ellipses. The real axis, that is, the rays $\theta = 0$ and $\theta = \pi$, are mapped onto the part of the real axis from $w = 2$ via ∞ to $w = -2$. The y -axis is mapped onto the v -axis. Any other pair of rays $\theta = \theta_0$ and $\theta = \theta_0 + \pi$ is mapped onto the two branches of the same hyperbola.

An exterior region of one of the above ellipses is free of branch points and corresponds to either the interior or the exterior of the corresponding circle in the z -plane, depending on the sheet of the Riemann surface to which the region belongs. In particular the full w -plane corresponds to the interior or the exterior of the unit circle $|z| = 1$, as was mentioned before.

The mapping (5) transforms suitable circles into airfoils with a sharp trailing edge whose interior angle is zero; these airfoils are known as Joukowski¹ airfoils. Since the airfoil to be obtained has a sharp edge, it is clear that the circle to be mapped must pass through one of the points $z = \pm 1$ where the mapping is not conformal. Let us choose a circle C through $z = -1$ and such that $z = 1$ lies inside C . The simplest way of determining the image is

¹ NIKOLAI JEGOROVICH JOUKOWSKI (1847–1921), Russian mathematician.

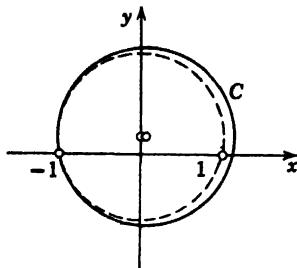


Fig. 307. Joukowski airfoil.

the graphical vector addition of the vectors corresponding to z and $1/z$ (where the latter can be obtained from the former as shown in Sec. 11.3). This leads to the result shown in Fig. 307. Further details are included in Ref. [G7] in Appendix 1.

A more elaborate treatment of Riemann surfaces can be found in Ref. [G8].

Problems

1. Consider $w = \sqrt{z}$. Find the path of the image point w of a point z which moves twice around the unit circle, starting from the initial position $z = 1$.
2. Show that the Riemann surface of $w = \sqrt[3]{z}$ consists of three sheets and has a branch point of third order at $z = 0$. Find the path of the image point w of a point z which moves three times around the unit circle starting from the initial position $z = 1$.
3. Consider the Riemann surfaces of $w = \sqrt[4]{z}$ and $w = \sqrt[5]{z}$ in a fashion similar to that in Prob. 2.
4. Determine the path of the image of a point z under the mapping $w = \ln z$ as z moves several times around the unit circle.
5. Show that the Riemann surface of $w = \sqrt{(z-1)(z-4)}$ has branch points at $z = 1$ and $z = 4$ and consists of two sheets which may be cut along the line segment from 1 to 4 and joined crosswise. Hint: introduce polar coordinates $z-1 = r_1 e^{i\theta_1}$, $z-4 = r_2 e^{i\theta_2}$.
6. Show that the Riemann surface of $w = \sqrt{(1-z^2)(4-z^2)}$ has four branch points and two sheets which may be joined crosswise along the segments $-2 \leq x \leq -1$ and $1 \leq x \leq 2$ of the x -axis.

Determine the location of the branch points and the number of sheets of the Riemann surfaces of the following functions.

$$7. w = i\sqrt{z} \quad 8. w = \sqrt{z-i} \quad 9. w = \sqrt[3]{z-i} \quad 10. w = \sqrt{z^2-1}$$

Chapter 12

Complex Integrals

There are two main reasons why integrals in the complex plane are so important. The practical reason is that in applications there occur real integrals which can be evaluated by complex integration, while the usual methods of real integral calculus are not successful. The other reason is of a theoretical nature and results from the fact that some basic properties of analytic functions (such as the existence of higher derivatives) cannot be proved by other methods. In other words, questions which superficially concern only the complex *differential* calculus must be considered by means of *integration*. This fact indicates basic differences between real and complex calculus.

In this chapter we shall first define complex integrals. The most fundamental result will be Cauchy's integral theorem (Sec. 12.3) from which the important Cauchy integral formula (Sec. 12.5) will follow. In Sec. 12.6 we shall prove that, if a function is analytic, it has derivatives of any order. This means that in this respect complex analytic functions behave much more simply than real functions of a real variable.

(Integration by means of residues and applications to real integrals will be considered in Chap. 15.)

Prerequisite for this chapter: Chap. 10.

References: Appendix 1, Part G.

Answers to problems: Appendix 2.

✓ 12.1 *Line Integral in the Complex Plane*

As in real calculus, we distinguish between definite integrals and indefinite integrals or antiderivatives. An **indefinite integral** is a function whose derivative equals a given analytic function in a region. By inversion of known differentiation formulas we may find many types of indefinite integrals.

Let us now define *definite integrals*, or *line integrals*, of a complex function $f(z)$ where $z = x + iy$. We shall see that this definition will be a natural

generalization of the familiar definition of a real definite integral, and the consideration will be similar to that in Sec. 6.1. In the case of a definite integral the path of integration is an interval of the real axis. In the case of a complex definite integral we shall integrate along a curve¹ in the complex plane.

Let C be a smooth curve (cf. Sec. 11.2) in the complex z -plane. Then we may represent C in the form

$$(1) \quad z = z(t) = x(t) + iy(t) \quad (a \leq t \leq b)$$

where $z(t)$ has a continuous derivative $\dot{z}(t) \neq 0$ for all t , so that C is rectifiable (cf. Sec. 5.12) and has a unique tangent at each point. We remember that the positive direction along C corresponds to the sense of increasing values of the parameter t .

Let $f(z)$ be a single-valued and continuous function which is defined (at least) at each point of C . We subdivide C by points $z_0, z_1, \dots, z_{n-1}, z_n (= Z)$, cf. Fig. 308. On each portion of subdivision we choose an arbitrary point, say, a point ζ_1 between z_0 and z_1 , a point ζ_2 between z_1 and z_2 , etc., and form the sum

$$(2) \quad S_n = \sum_{m=1}^n f(\zeta_m) \Delta z_m$$

where

$$\Delta z_m = z_m - z_{m-1}.$$

This we do for each $n = 2, 3, \dots$ in a completely independent manner, but such that the greatest $|\Delta z_m|$ approaches zero as n approaches infinity. In this way we obtain a sequence of complex numbers S_2, S_3, \dots . The limit of this sequence is called the **line integral** (or simply the integral) of $f(z)$ along the oriented curve C and is denoted by

$$(3) \quad \int_C f(z) dz.$$

The curve C is called the **path of integration**.

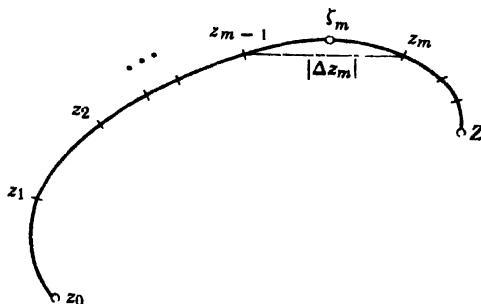


Fig. 308. Complex line integral.

¹ Actually along a portion, or arc, of a curve. For the sake of simplicity we shall use the single term "curve" to denote an entire curve as well as a portion of it.

Throughout the following considerations all paths of integration for complex line integrals will be assumed to be piecewise smooth, that is, will consist of finitely many smooth curves.

From our assumptions the existence of the line integral (3) follows. In fact, let $f(z) = u(x, y) + iv(x, y)$ and set

$$\zeta_m = \xi_m + i\eta_m \quad \text{and} \quad \Delta z_m = \Delta x_m + i\Delta y_m.$$

Then (2) may be written

$$(4) \quad S_n = \sum (u + iv)(\Delta x_m + i\Delta y_m)$$

where $u = u(\xi_m, \eta_m)$, $v = v(\xi_m, \eta_m)$ and we sum over m from 1 to n . We may now split up S_n into four sums:

$$S_n = \sum u \Delta x_m - \sum v \Delta y_m + i[\sum u \Delta y_m + \sum v \Delta x_m].$$

These sums are real. Since f is continuous, u and v are continuous. Hence, if we let n approach infinity in the aforementioned way, then the greatest Δx_m and Δy_m will approach zero and each sum on the right becomes a real line integral:

$$(5) \quad \lim_{n \rightarrow \infty} S_n = \int_C f(z) dz = \int_C u dx - \int_C v dy + i \left[\int_C u dy + \int_C v dx \right].$$

This shows that the line integral (3) exists and its value is independent of the choice of subdivisions and intermediate points ζ_m .

Furthermore, as in Sec. 6.2 we may convert each of those real line integrals to a definite integral by the use of the representation (1) of the curve C :

$$(6) \quad \int_C f(z) dz = \int_a^b u \dot{x} dt - \int_a^b v \dot{y} dt + i \left[\int_a^b u \dot{y} dt + \int_a^b v \dot{x} dt \right]$$

where $u = u[x(t), y(t)]$, $v = v[x(t), y(t)]$, and dots denote derivatives with respect to t .

Without causing misinterpretation, we may also write

$$\int_C f(z) dz = \int_a^b (u + iv)(\dot{x} + i\dot{y}) dt$$

or more briefly

$$(6^*) \quad \int_C f(z) dz = \int_a^b f[z(t)] \dot{z}(t) dt.$$

Let us consider some basic examples.

Example 1. Integrate $f(z) = 1/z$ once around the unit circle C in the counterclockwise sense, starting from $z = 1$. We may represent C in the form

$$(7) \quad z(t) = \cos t + i \sin t \quad (0 \leq t \leq 2\pi).$$

Then we have

$$\dot{z}(t) = -\sin t + i \cos t,$$

and by (6*) the integral under consideration becomes

$$\int_C \frac{dz}{z} = \int_0^{2\pi} \frac{1}{\cos t + i \sin t} (-\sin t + i \cos t) dt = i \int_0^{2\pi} dt = 2\pi i,$$

a fundamental result which we shall use in our further consideration.

Clearly, instead of (7) we may write more simply

$$(7') \quad z(t) = e^{it} \quad (0 \leq t \leq 2\pi)$$

Then we obtain by differentiation

$$\dot{z}(t) = ie^{it}, \quad dz = ie^{it} dt,$$

and from this the same result as before:

$$(8) \quad \int_C \frac{dz}{z} = \int_0^{2\pi} \frac{1}{e^{it}} ie^{it} dt = i \int_0^{2\pi} dt = 2\pi i.$$

Example 2. Integrate $f(z) = \operatorname{Re} z = x$ along the segment from $z_0 = 0$ to $Z = 1 + i$ (path C_1 in Fig. 309).

The segment may be represented in the form

$$z(t) = x(t) + iy(t) = (1+i)t \quad (0 \leq t \leq 1).$$

Then

$$f[z(t)] = \operatorname{Re} z(t) = x(t) = t, \quad dz = (1+i) dt.$$

Therefore, we obtain the result

$$\int_{C_1} \operatorname{Re} z dz = \int_0^1 t(1+i) dt = (1+i) \int_0^1 t dt = \frac{1}{2}(1+i).$$

Let us now integrate $f(z) = \operatorname{Re} z = x$ along the real axis from 0 to 1, and then vertically to $1 + i$ (path C_2 in Fig. 309). We may represent the first part of C_2 by

$$z = z(t) = t \quad (0 \leq t \leq 1)$$

and the last part in the form

$$z(t) = 1 + i(t-1) \quad (1 \leq t \leq 2).$$

Then the entire path corresponds to the interval $0 \leq t \leq 2$. On the first part, $\operatorname{Re} z = t$, $dz = dt$, and on the last part, $\operatorname{Re} z = 1$, $dz = i dt$. Therefore,

$$\int_{C_2} \operatorname{Re} z dz = \int_0^1 t dt + \int_1^2 i dt = \frac{1}{2} + i.$$

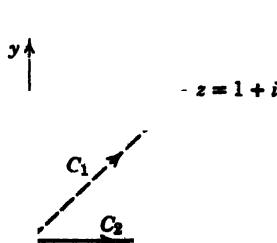


Fig. 309. Paths in Example 2.

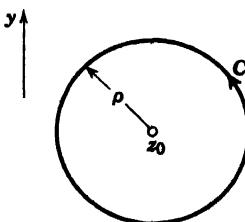


Fig. 310. Path in Example 3.

We note that the last part of C_2 could be represented equally well in the form

$$z(t) = 1 + it \quad (0 \leq t \leq 1);$$

then the limits of the last integral are 0 and 1, and its value is the same as before.

By comparing our two results we see that in the case of the function $\operatorname{Re} z$ (which is not analytic) the value of the integral depends not only on the end points of the path, but also on its geometric shape.

Example 3. Let $f(z) = (z - z_0)^m$ where m is an integer and z_0 is a constant. Integrate counterclockwise around the circle C of radius ρ with center at z_0 (cf. Fig. 310 on p. 591).

We may represent C in the form

$$z(t) = z_0 + \rho(\cos t + i \sin t) = z_0 + \rho e^{it} \quad (0 \leq t \leq 2\pi).$$

Then we have

$$(z - z_0)^m = \rho^m e^{imt}, \quad dz = i\rho e^{it} dt,$$

and we obtain

$$\int_C (z - z_0)^m dz = \int_0^{2\pi} \rho^m e^{imt} i\rho e^{it} dt = i\rho^{m+1} \int_0^{2\pi} e^{i(m+1)t} dt$$

The case $m = -1$ was considered in Ex. 2, and when $m \neq -1$, we obtain [cf. (14) in Sec. 10.5]

$$\int_0^{2\pi} e^{i(m+1)t} dt = \left[\frac{e^{i(m+1)t}}{i(m+1)} \right]_0^{2\pi} = 0 \quad (m \neq -1).$$

The result is

$$(9) \quad \int_C (z - z_0)^m dz = \begin{cases} 2\pi i & (m = -1), \\ 0 & (m \neq -1 \text{ and integral}). \end{cases}$$

Example 4. Let $f(z) = k = \text{const}$, and C be any curve joining two points z_0 and Z . In this case we may use the definition of the line integral as the limit of the sum S_n given by (2). We have

$$S_n = \sum_{m=1}^n k \Delta z_m = k[(z_1 - z_0) + (z_2 - z_1) + \cdots + (Z - z_{n-1})] = k(Z - z_0).$$

From this we immediately have the result

$$\int_C k dz = \lim_{n \rightarrow \infty} S_n = k(Z - z_0).$$

We note that the value of this integral depends only on z_0 and Z , not on the geometric shape of the path joining these points. If in particular C is closed, then $z_0 = Z$ and the integral is zero.

Example 5. Let $f(z) = z$ and C be any curve joining two points z_0 and Z . We again use (2). Taking $\zeta_m = z_m$ we obtain

$$S_n = \sum_{m=1}^n z_m \Delta z_m = z_1(z_1 - z_0) + z_2(z_2 - z_1) + \cdots + Z(Z - z_{n-1}).$$

Taking $\zeta_m = z_{m-1}$ we find

$$S_n^* = \sum_{m=1}^n z_{m-1} \Delta z_m = z_0(z_1 - z_0) + z_1(z_2 - z_1) + \cdots + z_{n-1}(Z - z_{n-1}).$$

Addition of these two sums yields $S_n + S_n^* = Z^2 - z_0^2$, as can easily be verified. From this we have

$$\lim_{n \rightarrow \infty} (S_n + S_n^*) = 2 \int_{z_0}^Z z dz = Z^2 - z_0^2.$$

Hence,

$$\int_z^Z z \, dz = \frac{1}{2}(Z^2 - z_0^2)$$

along every path joining z_0 and Z . If, in particular, C is a closed path, then $z_0 = Z$ and

$$(10) \quad \int_C z \, dz = 0.$$

Note that this result follows also from Green's theorem (cf. Sec. 6.4) by the use of formula (6) in this section.

Problems

Represent the line segments from A to B in the form $z = z(t)$ where

- | | |
|------------------------------------|----------------------------------|
| 1. $A: z = 0, B: z = 1 + 2i$ | 2. $A: z = 0, B: z = 4 - 8i$ |
| 3. $A: z = 1 + i, B: z = 3 - 4i$ | 4. $A: z = 3i, B: z = 4 - i$ |
| 5. $A: z = -2 + i, B: z = -2 + 4i$ | 6. $A: z = 3 - i, B: z = 5 - 7i$ |

Represent the following curves in the form $z = z(t)$.

7. The circle of radius 5 with center at $z = 0$
8. The circle of radius 3 with center at $z = 1 - 2i$
9. The parabola $y = x^2$ from $(0, 0)$ to $(3, 9)$
10. The parabola $y = 2x^2 - 1$ from $(2, 7)$ to $(3, 17)$
11. The ellipse $x^2 + 9y^2 = 9$
12. The ellipse $4(x - 1)^2 + 9(y + 2)^2 = 36$
13. The hyperbola $y = 1/x$ from $(1, 1)$ to $(4, 1/4)$
14. The hyperbola $y = -1 + 3/x$ from $(1, 2)$ to $(3, 0)$

What curves are represented by the following functions?

- | | |
|---|---|
| 15. $1 + (2 - t)t, \quad 0 \leq t \leq 1$ | 16. $i - t + 3it, \quad -1 \leq t \leq 2$ |
| 17. $3i + 3e^{it}, \quad 0 \leq t \leq \pi$ | 18. $4 + 2e^{it}, \quad -\pi \leq t \leq 0$ |
| 19. $t + 3t^2i, \quad -1 \leq t \leq 2$ | 20. $i + t - t^3i, \quad 0 \leq t \leq 3$ |

Evaluate $\int_C 3z^2 \, dz$, where C is

21. The line segment from 0 to $2i$
22. The line segment from 0 to $1 + i$
23. The line segment in Prob. 3
24. The line segment in Prob. 4
25. The parabola in Prob. 9
26. The unit circle (cf. Sec. 10.1)
27. The boundary of the triangle with vertices $0, 1, i$
28. The boundary of the square with vertices $0, 2, 2 + 2i, 2i$
- 29. Evaluate $\int_C |z| \, dz$ from $A: z = -i$ to $B: z = i$ along (a) the line segment AB , (b) the unit circle in the left half plane, (c) the unit circle in the right half plane.
30. Evaluate $\int_C (1/\sqrt{z}) \, dz$ from 1 to -1 (a) along the upper semicircle $|z| = 1$, (b) along the lower semicircle $|z| = 1$, where \sqrt{z} is the principal value of the square root.

12.2 Basic Properties of the Complex Line Integral

From the definition of a complex line integral as the limit of a sum we may immediately obtain the following properties.

If we decompose the path C into two portions C_1 and C_2 (Fig. 311), then

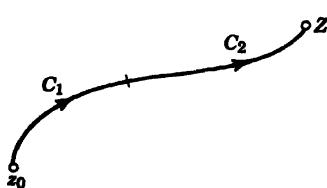


Fig. 311. Formula (1).

$$(1) \quad \int_C f(z) dz = \int_{C_1} f(z) dz + \int_{C_2} f(z) dz.$$

If we reverse the sense of integration, the sign of the value of the integral changes:

$$(2) \quad \int_{z_0}^Z f(z) dz = - \int_Z^{z_0} f(z) dz;$$

here the path C with end points z_0 and Z is the same; on the left we integrate from z_0 to Z , and on the right in the opposite sense.

A sum of two (or more) functions may be integrated term by term, and constants may be taken out from under the integral sign, that is,

$$(3) \quad \int_C [k_1 f_1(z) + k_2 f_2(z)] dz = k_1 \int_C f_1(z) dz + k_2 \int_C f_2(z) dz.$$

There will be a frequent necessity for estimating the absolute value of complex line integrals. The basic formula is

$$(4) \quad \left| \int_C f(z) dz \right| \leq Ml$$

where l is the length of the path C and M is a real constant such that $|f(z)| \leq M$ everywhere on C .

To prove (4), we apply (3), Sec. 10.1, to the sum S_n defined by (2), Sec. 12.1, finding

$$|S_n| = \left| \sum_{m=1}^n f(\zeta_m) \Delta z_m \right| \leq \sum_{m=1}^n |f(\zeta_m)| |\Delta z_m| \leq M \sum_{m=1}^n |\Delta z_m|.$$

Now $|\Delta z_m|$ is the length of the chord whose end points are z_{m-1} and z_m , cf. Fig. 308 in Sec. 12.1. The sum on the right thus represents the length L of the broken line of chords whose end points are $z_0, z_1, \dots, z_n (= Z)$. If n approaches infinity such that the greatest $|\Delta z_m|$ approaches zero, then L approaches the length l of the curve C , by definition of the length of a curve (cf. Sec. 5.12). From this, formula (4) follows.

Example 1. Let $f(z) = 1/z$ and integrate once around the circle $|z| = \rho$. Then $l = 2\pi\rho$ and $|f(z)| = 1/\rho$ on the circle. Hence, by (4),

$$\left| \int_C \frac{dz}{z} \right| \leq \frac{1}{\rho} 2\pi\rho = 2\pi \quad (\text{cf. Ex. 1, Sec. 12.1}).$$

Example 2. The path C_2 in Ex. 2, Sec. 12.1, has the length $l = 2$, and $|\operatorname{Re} z| \leq 1$ on C_2 . Therefore, by (4),

$$\left| \int_{C_2} \operatorname{Re} z \, dz \right| \leq 2.$$

Problems

1. Verify (1) for $f(z) = 1/z$ where C is the unit circle, C_1 its upper half, and C_2 its lower half.
2. Verify (2) for $f(z) = z^2$ where C is the line segment from $-1 - i$ to $1 + i$.
3. Verify (3) for $k_1 f_1 + k_2 f_2 = 3z - z^2$ where C is the upper half of the unit circle from 1 to -1 .
4. Verify (4) for $f(z) = 1/z$ and C as in Prob. 3.

Evaluate $\int_C f(z) \, dz$ where

5. $f(z) = az + b$, C as in Prob. 2
6. $f(z) = z^3 + 2z^{-1}$, C the unit circle (counterclockwise)
7. $f(z) = z^2 + 3z^{-4}$, C as in Prob. 3
8. $f(z) = 2z - z^{-1} + 2z^{-2}$, C the unit circle (counterclockwise)
9. $f(z) = e^z$, C the line segment from 0 to $1 + \pi i/2$
10. $f(z) = e^z$, C the unit circle (clockwise)
11. $f(z) = \cos z$, C the line segment from πi to $2\pi i$
12. $f(z) = \sin z$, C as in Prob. 11
13. $f(z) = \sin z$, C the line segment from 0 to i
14. $f(z) = \sinh z$, C as in Prob. 13
15. $f(z) = \cosh z$, C as in Prob. 13

Using (4), find upper bounds for the following integrals where C is the line segment from 0 to $2 + 2i$.

$$16. \int_C z \, dz \quad 17. \int_C z^2 \, dz \quad 18. \int_C e^z \, dz \quad 19. \int_C (3z + 1) \, dz \quad 20. \int_C \ln(z + 1) \, dz$$

12.3 Cauchy's Integral Theorem

Cauchy's integral theorem is very important in complex analysis and has various theoretical and practical consequences. To state this theorem, we shall need the following concept.

The interior of a simple closed curve (that is, a curve which is closed and has no self-intersections) is called a **simply connected domain**. A domain which is not simply connected is said to be *multiply connected*. For example, the interior of a circle ("circular disk"), ellipse, square, etc., is simply connected and a circular ring (cf. Sec. 10.1) is multiply connected (more precisely: doubly² connected).

² A domain is said to be *p-fold connected* if its boundary consists of p parts without common points. For the annulus, $p = 2$, because the boundary consists of two circles having no points in common.

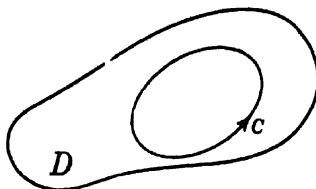


Fig. 312. Cauchy's integral theorem.

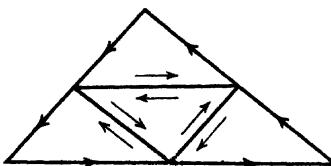


Fig. 313. Proof of Cauchy's integral theorem.

Cauchy's integral theorem. If $f(z)$ is analytic³ in a simply connected bounded domain D , then

$$(1) \quad \int_C f(z) dz = 0$$

on every simple closed path C in D .

Cauchy's proof. From (5), Sec. 12.1, we have

$$\int_C f(z) dz = \int_C (u dx - v dy) + i \int_C (u dy + v dx).$$

$f(z)$ is analytic and, therefore $f'(z)$ exists. Cauchy made the *additional assumption* that $f'(z)$ is continuous. Then u and v have continuous first partial derivatives in D , as follows from (3) and (4), Sec. 10.3. Green's theorem (cf. Sec. 6.4) (with u and $-v$ instead of f and g) is applicable, and

$$\int_C (u dx - v dy) = \iint_R \left(-\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) dx dy$$

where R is the region bounded by C . The second Cauchy-Riemann equation (Sec. 10.3) shows that the integrand on the right is identically zero. Hence, the integral on the left is zero. In the same fashion it follows by the use of the first Cauchy-Riemann equation that the last integral in the above formula is zero. This completes Cauchy's proof.

Goursat's⁴ proof. Goursat proved Cauchy's theorem without assuming that $f'(z)$ is continuous. This progress is quite important. We start with the case when C is the boundary of a triangle. By joining the midpoints of the sides of the triangle we subdivide it into four congruent triangles (Fig. 313). Then

$$\int_C f dz = \int_{C_1} f dz + \int_{C_{II}} f dz + \int_{C_{III}} f dz + \int_{C_{IV}} f dz$$

where C_1, \dots, C_{IV} are the boundaries of the triangles; in fact, on the right side we integrate along each of the three segments of subdivision in both

³ Remember our general assumption that the functions under consideration are single-valued. Actually, if a function is analytic in a *simply connected* domain D , it is single-valued in D . The proof of this fact is rather difficult.

⁴ EDOUARD GOURSAT (1858–1936), French mathematician.

possible directions, the corresponding integrals cancel out in pairs, and the sum of the integrals on the right equals the integral on the left. Among the four integrals on the right there must be one, call its path C_1 , for which

$$\left| \int_{C_1} f dz \right| \leq 4 \left| \int_{C_1} f dz \right|$$

because not each of the absolute values of the four integrals can be smaller than $\frac{1}{4}$ of the absolute value of the sum of those integrals. This follows readily from (3) in Sec. 10.1.

We subdivide the triangle bounded by C_1 as before and select a triangle of subdivision with boundary C_2 for which

$$\left| \int_{C_2} f dz \right| \leq 4 \left| \int_{C_2} f dz \right|, \quad \text{and then } \left| \int_C f dz \right| \leq 4^2 \left| \int_{C_2} f dz \right|$$

Continuing in this fashion, we obtain a sequence of triangles T_1, T_2, \dots with boundaries C_1, C_2, \dots which are similar and such that T_n lies in T_m when $n > m$, and

$$(2) \quad \left| \int_C f dz \right| \leq 4^n \left| \int_{C_n} f dz \right|, \quad n = 1, 2,$$

Let z_0 be the point which belongs to all these triangles. Since f is differentiable at $z = z_0$, the derivative $f'(z_0)$ exists, and we may write

$$(3) \quad f(z) = f(z_0) + (z - z_0)f'(z_0) + h(z)(z - z_0).$$

By integrating over the boundary C_n of the triangle T_n we have

$$\int_{C_n} f(z) dz = \int_{C_n} f(z_0) dz + \int_{C_n} (z - z_0)f'(z_0) dz + \int_{C_n} h(z)(z - z_0) dz.$$

Since $f(z_0)$ and $f'(z_0)$ are constants, it follows from the results in Exs. 4 and 5, Sec. 12.1, that the first two integrals on the right are zero. Hence,

$$\int_{C_n} f(z) dz = \int_{C_n} h(z)(z - z_0) dz.$$

By dividing (3) by $z - z_0$, transposing two terms to the left, and taking absolute values we obtain

$$\left| \frac{f(z) - f(z_0)}{z - z_0} - f'(z_0) \right| = |h(z)|.$$

From this and (5) in Sec. 10.2 we see that for given $\epsilon > 0$ we can find a $\delta > 0$ such that

$$|h(z)| < \epsilon \quad \text{when} \quad |z - z_0| < \delta.$$

We may now take n so large that the triangle T_n lies in the disk $|z - z_0| < \delta$. Let l_n be the length of C_n . Then $|z - z_0| \leq l_n/2$ for all z on C_n and z_0 in T_n . By applying (4) in Sec. 12.2 we thus obtain

$$(4) \quad \left| \int_{C_n} f(z) dz \right| = \left| \int_{C_n} h(z)(z - z_0) dz \right| < \epsilon \frac{l_n}{2} l_n = \frac{\epsilon}{2} l_n^2.$$

Let l be the length of C . Then the path C_1 has the length $l_1 = l/2$, the path C_2 has the length $l_2 = l_1/2 = l/4$, etc., and C_n has the length

$$l_n = \frac{l}{2^n}.$$

From (2) and (4) we thus obtain

$$\left| \int_C f dz \right| \leq 4^n \left| \int_{C_n} f dz \right| < 4^n \frac{\epsilon}{2} l_n^2 = 4^n \frac{\epsilon}{2} \frac{l^2}{4^n} = \frac{\epsilon}{2} l^2.$$

By choosing $\epsilon (>0)$ sufficiently small we can make the expression on the right as small as we please, while the expression on the left is the definite value of an integral. From this we conclude that this value must be zero, and the proof is complete.

The proof for the case in which C is the boundary of a polygon follows from the previous proof by subdividing the polygon into triangles (Fig. 314). The integral corresponding to each such triangle is zero. The sum of these integrals is equal to the integral over C , because we integrate along each segment of subdivision in both directions, the corresponding integrals cancel out in pairs, and we are left with the integral over C .

The case of a general simple closed path C can be reduced to the preceding one by inscribing in C a closed polygon P of chords, which approximates C "sufficiently accurately," and it can be shown that one can find a polygon P such that the integral over P differs from that over C by less than any preassigned positive real number ϵ , no matter how small. The details of this proof are somewhat involved and can be found in Ref. [G5] in Appendix 1.

Example 1.

$$\int_C e^z dz = 0$$

for any closed path, because e^z is analytic for all z

Example 2.

$$\int_C \frac{dz}{z^2} = 0,$$

where C is the unit circle (cf. Sec. 12.1). This result does not follow from Cauchy's theorem, because $f(z) = 1/z^2$ is not analytic at $z = 0$. It illustrates that the condition that f be analytic in D is sufficient rather than necessary for (1) to be true.

Example 3.

$$\int_C \frac{dz}{z} = 2\pi i,$$

the integration being taken around the unit circle in the counterclockwise sense (cf. Sec. 12.1). C lies in the annulus $1/2 < |z| < 3/2$ where $1/z$ is analytic, but this domain is not simply connected, so that Cauchy's theorem cannot be applied. Hence the condition that the domain D be simply connected is quite essential.

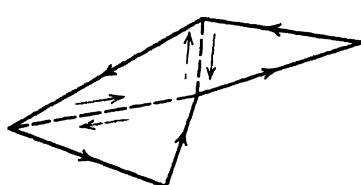


Fig. 314. Proof of Cauchy's integral theorem for a polygon.

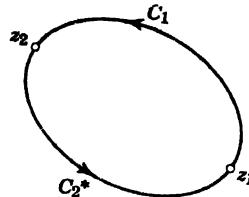


Fig. 315. Formula (5').

If we subdivide the path C in Cauchy's theorem into two arcs C_1 and C_2^* (Fig. 315), then (1) takes the form

$$\int_C f \, dz = \int_{C_1} f \, dz + \int_{C_2^*} f \, dz = 0.$$

Consequently,

$$(5') \quad -\int_{C_2^*} f \, dz = \int_{C_1} f \, dz.$$

If we reverse the sense of integration along C_2^* , then the integral over C_2^* is multiplied by -1 , and we obtain (cf. Fig. 316)

$$(5) \quad \int_{C_2} f(z) \, dz = \int_{C_1} f(z) \, dz.$$

Hence, if f is analytic in D , and C_1 and C_2 are any paths in D joining two points in D and having no further points in common, then (5) holds.

If those paths C_1 and C_2 have finitely many points in common (Fig. 317), then (5) continues to hold; this follows by applying the previous result to the portions of C_1 and C_2 between each pair of consecutive points of intersection.

It is even true that (5) holds for any paths entirely in the domain D where $f(z)$ is analytic and joining any points z_1 and z_2 in D . Then we say that the integral of $f(z)$ from z_1 to z_2 is independent of path in D . (Of course, the value of the integral depends on the choice of z_1 and z_2 .)

The proof would require an additional consideration of the case in which C_1 and C_2 have infinitely many points of intersection, and will not be presented here.

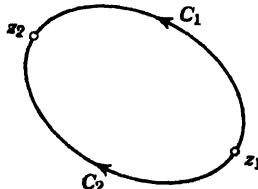


Fig. 316. Formula (5).

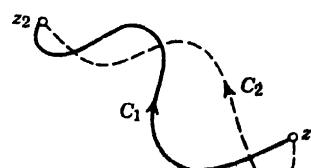


Fig. 317. Paths having finitely many intersections.

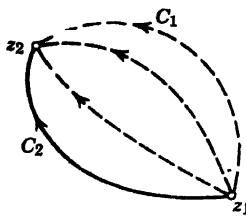


Fig. 318. Continuous deformation of path.

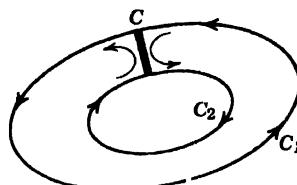


Fig. 319. Doubly connected domain.

We may imagine that the path C_2 in (5) was obtained from C_1 by a continuous deformation (Fig. 318). It follows that in a given integral we may impose a continuous deformation on the path of integration (keeping the end points fixed); as long as we do not pass through a point where $f(z)$ is not analytic, the value of the line integral will not change under such a deformation. This is often called the principle of deformation of path.

A multiply connected domain can be cut such that the resulting domain becomes simply connected. For a doubly connected domain D^* we need one cut \tilde{C} (Fig. 319). If $f(z)$ is analytic in D^* and at each point of C_1 and C_2 then, since C_1 , C_2 , and \tilde{C} bound a simply connected region, it follows from Cauchy's theorem that the integral of f taken over C_1 , \tilde{C} , C_2 in the sense indicated by the arrows in Fig. 319 has the value zero. Since we integrate along \tilde{C} in both directions, the corresponding integrals cancel out, and we obtain

$$(6) \quad \int_{C_1} f(z) dz + \int_{C_2} f(z) dz = 0$$

where one of the curves is traversed in the counterclockwise sense and the other in the opposite sense.

Equation (6) may also be written

$$(7) \quad \int_{C_1} f(z) dz = \int_{C_2} f(z) dz,$$

where C_1 and C_2 are now traversed in the same direction (Fig. 320). We remember that (7) holds under the assumption that $f(z)$ is analytic in the domain bounded by C_1 and C_2 and at each point of C_1 and C_2 .

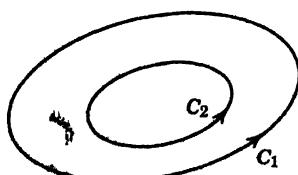


Fig. 320. Paths in (7).

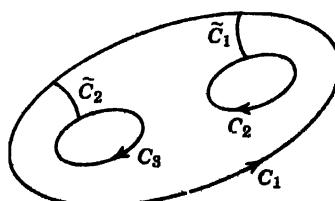


Fig. 321. Triply connected domain.

For more complicated domains we may need more than one cut, but the basic idea remains the same as before. For instance, for the domain in Fig. 321,

$$\int_{C_1} f(z) dz + \int_{C_2} f(z) dz + \int_{C_3} f(z) dz = 0$$

where C_2 and C_3 are traversed in the same sense and C_1 is traversed in the opposite sense.

We mention that a simple closed path is sometimes called a *contour*, and integrals along such a path are known as **contour integrals**.

Example 4. Let C_1 be the unit circle $|z| = 1$ and C_2 the circle $|z| = 1/2$. Then, by (7),

$$\int_{C_2} \frac{dz}{z} = \int_{C_1} \frac{dz}{z} = 2\pi i \quad (\text{cf. Ex. 1 in Sec. 12.1}),$$

where both paths are traversed in the counterclockwise sense.

Example 5. From Ex. 3 in Sec. 12.1 and the principle of deformation of path it follows that

$$\int_C (z - z_0)^m dz = \begin{cases} 2\pi i & (m = -1) \\ 0 & (m \neq -1 \text{ and integral}) \end{cases}$$

where C is any contour containing the point z_0 in its interior and the integration is taken around C in the counterclockwise sense.

Problems

- Verify Cauchy's theorem for $\int_C z^2 dz$ where C is the boundary of the triangle with vertices 0, 2, and $2i$.
- Show that the integral of $1/z^3$ around the unit circle is zero. Does this follow from Cauchy's theorem?
- For what simple closed paths is the integral of $1/z$ equal to zero?

To which of the following integrals does Cauchy's theorem apply?

- $\int_C \frac{dz}{1+z^2}$, $C: |z+1|=1$
- $\int_C \frac{dz}{1+z^2}$, $C: |z-1|=2$
- $\int_C \frac{\sin z}{z} dz$, $C: |z|=1$
- $\int_C \frac{\cos z}{z} dz$, $C: |z|=1$

- Using the principle of deformation of path and

$$\frac{2z-1}{z^2-z} = \frac{1}{z} + \frac{1}{z-1}, \quad \text{show that} \quad \int_C \frac{2z-1}{z^2-z} dz = \int_C \frac{dz}{z} + \int_C \frac{dz}{z-1} = 4\pi i$$

where C is the path in Fig. 322.

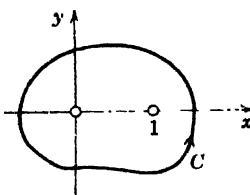


Fig. 322. Path in Problem 8.

9. Integrate z^2 from 0 to $1 + i$ along C_1 , C_2 , and C_3 in Fig. 323.

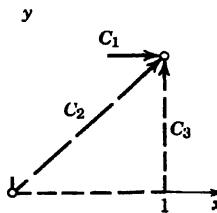


Fig. 323. Problem 9.

Evaluate the following integrals. (*Hint:* if necessary, represent the integrand in terms of partial fractions.)

10. $\int_C \frac{dz}{z}$, $C: |z - 2| = 1$ (clockwise)

11. $\int_C \frac{z^2 - z + 1}{z^3 - z^2} dz$, $C: (a) |z| = 2, (b) |z| = 1/2$ (counterclockwise)

12. $\int_C \frac{dz}{z^2 - 1}$, $C: (a) |z| = 2, (b) |z - 1| = 1$ (clockwise)

13. $\int_C \frac{z}{z^2 + 1} dz$, $C: (a) |z| = 2, (b) |z + i| = 1$ (counterclockwise)

14. $\int_C \frac{dz}{z^2 + 1}$, $C: (a) |z + i| = 1, (b) |z - i| = 1$ (counterclockwise)

15. $\int_C \frac{e^z}{z} dz$, C consists of the circles $|z| = 2$ (counterclockwise) and $|z| = 1$ (clockwise)

16. $\int_C \frac{\cos z}{z^2} dz$, $C: |z - 2i| = 1$ (counterclockwise)

17. $\int_C \frac{3z + 1}{z^3 - z} dz$, $C: (a) |z| = 1/2, (b) |z| = 2$ (counterclockwise)

18. $\int_C \frac{2z^3 + z^2 + 4}{z^4 + 4z^2} dz$, $C: |z - 2| = 4$ (clockwise)

19. $\int_C \frac{dz}{z^4 + 4z^2}$, C consists of $|z| = 3/2$ (clockwise) and $|z| = 1$ (counterclockwise)

20. $\int_C \cosh^2 z dz$, $C: |z| = 1$

12.4 Evaluation of Line Integrals by Indefinite Integration

Using Cauchy's integral theorem, we want to show that in many cases complex line integrals can be evaluated by a very simple method, namely, indefinite integration.

Suppose that $f(z)$ is analytic in a simply connected domain D , and let z_0 be any fixed point in D . Then the integral

$$\int_{z_0}^z f(z^*) dz^*$$

is a function of z for all paths which lie in D and join z_0 and z , and we may write

$$(1) \quad F(z) = \int_{z_0}^z f(z^*) dz^*.$$

Let us prove that $F(z)$ is an analytic function of z in D , and $F'(z) = f(z)$.

We keep z fixed. Since D is a domain, a neighborhood N of z belongs to D . In N we choose a point $z + \Delta z$. Then the line segment with end points z and $z + \Delta z$ is in D , and from (1) we obtain

$$F(z + \Delta z) - F(z) = \int_{z_0}^{z + \Delta z} f(z^*) dz^* - \int_{z_0}^z f(z^*) dz^* = \int_z^{z + \Delta z} f(z^*) dz^*,$$

where we may integrate from z to $z + \Delta z$ along that segment (Fig. 324). Hence,

$$\frac{F(z + \Delta z) - F(z)}{\Delta z} - f(z) = \frac{1}{\Delta z} \int_z^{z + \Delta z} [f(z^*) - f(z)] dz^*$$

because z is kept fixed and, therefore,

$$-\frac{1}{\Delta z} \int_z^{z + \Delta z} f(z) dz^* = -\frac{f(z)}{\Delta z} \int_z^{z + \Delta z} dz^* = -f(z).$$

Now $f(z)$ is continuous. Hence, an $\epsilon > 0$ being given, we can find a number $\delta > 0$ such that

$$|f(z^*) - f(z)| < \epsilon \quad \text{when} \quad |z^* - z| < \delta.$$

Consequently, if $|\Delta z| < \delta$, then

$$\begin{aligned} \left| \frac{F(z + \Delta z) - F(z)}{\Delta z} - f(z) \right| &= \frac{1}{|\Delta z|} \left| \int_z^{z + \Delta z} [f(z^*) - f(z)] dz^* \right| \\ &< \frac{\epsilon}{|\Delta z|} \left| \int_z^{z + \Delta z} dz^* \right| = \epsilon; \end{aligned}$$

that is

$$(2) \quad F'(z) = \lim_{\Delta z \rightarrow 0} \frac{F(z + \Delta z) - F(z)}{\Delta z} = f(z).$$

From (1) it follows that if z_0 is replaced by another fixed point in D , the function $F(z)$ is changed by an additive constant. From (2) we see that $F(z)$

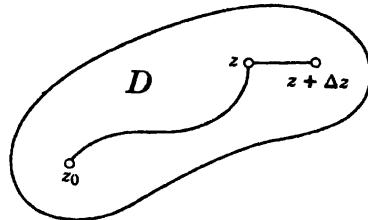


Fig. 324. Path of integration.

is an indefinite integral or antiderivative of $f(z)$, written

$$F(z) = \int f(z) dz;$$

that is, $F(z)$ is an analytic function in D whose derivative is $f(z)$.

If $F'(z) = f(z)$ and $G'(z) = f(z)$, then $F'(z) - G'(z) \equiv 0$ in D . Hence, the function $F(z) - G(z)$ is constant (cf. Prob. 17 at the end of Sec. 10.3). That is, the two indefinite integrals $F(z)$ and $G(z)$ differ only by a constant. In view of (1), we have for any points a and b in D and any path in D from a to b ,

$$\int_a^b f(z) dz = \int_{z_0}^b f(z) dz - \int_{z_0}^a f(z) dz = F(b) - F(a),$$

as in the case of real definite integrals, but it is important that the paths of integration lie in a simply connected domain D where $f(z)$ is analytic.

We may sum up our result as follows.

Theorem 1. If $f(z)$ is analytic in a simply connected domain D , and if $F(z)$ is an indefinite integral of $f(z)$, then for all paths in D joining two points a and b in D ,

$$(3) \quad \int_a^b f(z) dz = F(b) - F(a).$$

This theorem enables us to evaluate complex line integrals by means of indefinite integration.

Example 1.

$$\int_1^{1+4i} z^3 dz = \left[\frac{z^4}{4} \right]_1^{1+4i} = \frac{1}{3} [(1+4i)^4 - 1^4] = -\frac{47}{3} - 17i.$$

Example 2.

$$\int_1^{\pi/2} \cos z dz = \sin z \Big|_1^{\pi/2} = \sin \frac{\pi}{2} - \sin 1 = 1 - i \sinh 1.$$

Problems

Evaluate:

1. $\int_1^{2+i} z dz$

2. $\int_1^{3i} z^2 dz$

3. $\int_1^1 (z+1)^3 dz$

4. $\int_{1-i}^{1+i} z^3 dz$

5. $\int_1^{1+\pi i} e^z dz$

6. $\int_{\pi i}^{2\pi i} e^{2z} dz$

7. $\int_0^i ze^{z^2} dz$

8. $\int_{1-\pi i}^{1+\pi i} e^{z/2} dz$

9. $\int_0^{\pi i} \cos z dz$

10. $\int_0^{2\pi i} \sin 2z dz$

11. $\int_0^{\pi i} z \cos z^2 dz$

12. $\int_0^{\pi i} z \cos z dz$

13. $\int_{-\pi i}^{\pi i} \sin^2 z dz$

14. $\int_{1-i}^{1+i} \cos z dz$

15. $\int_0^{3i} \cosh z dz$

16. $\int_i^{1+i} \sinh z dz$

17. $\int_1^{2i} \sinh z dz$

18. $\int_1^{3i} z \sinh z^2 dz$

19. $\int_{-1}^i z \cosh z^2 dz$

20. $\int_{-\pi i}^{\pi i} z \cosh z dz$

✓ 12.5 Cauchy's Integral Formula

The most important consequence of Cauchy's integral theorem is Cauchy's integral formula. This formula and its conditions of validity may be stated as follows.

Theorem 1. *Let $f(z)$ be analytic and single-valued in a simply connected domain D . Then for any point z_0 in D and any simple closed path C in D which encloses z_0 (Fig. 325),*

$$(1) \quad \int_C \frac{f(z)}{z - z_0} dz = 2\pi i f(z_0) \quad (\text{Cauchy's integral formula}),$$

the integration being taken in the counterclockwise sense.

Proof. Writing $f(z) = f(z_0) + [f(z) - f(z_0)]$ and remembering that a constant may be taken out from under the integral sign, we have

$$(2) \quad \int_C \frac{f(z)}{z - z_0} dz = f(z_0) \int_C \frac{dz}{z - z_0} + \int_C \frac{f(z) - f(z_0)}{z - z_0} dz.$$

From Ex. 5 in Sec. 12.3 it follows that the first term on the right is equal to $2\pi i f(z_0)$. Hence we see that (1) holds, provided the last integral is zero. Now the integrand of this integral is analytic in D except at the point z_0 . We may thus replace C by a small circle K with center at z_0 , without changing the value of the integral (Fig. 326). Since $f(z)$ is analytic, it is continuous. Hence, an $\epsilon > 0$ being given, we can find a $\delta > 0$ such that

$$|f(z) - f(z_0)| < \epsilon \quad \text{for all } z \text{ in the disk } |z - z_0| < \delta.$$

Choosing the radius ρ of K smaller than δ , we thus have

$$\left| \frac{f(z) - f(z_0)}{z - z_0} \right| < \frac{\epsilon}{\rho}$$

at each point of K . The length of K is $2\pi\rho$. Hence, by (4) in Sec. 12.2,

$$\left| \int_K \frac{f(z) - f(z_0)}{z - z_0} dz \right| < \frac{\epsilon}{\rho} 2\pi\rho = 2\pi\epsilon.$$

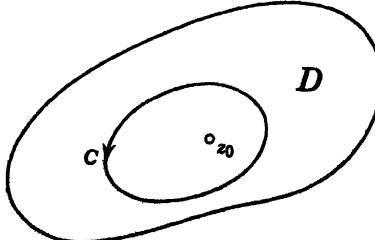


Fig. 325. Cauchy's integral formula. *

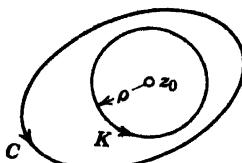


Fig. 326. Proof of Cauchy's integral formula.

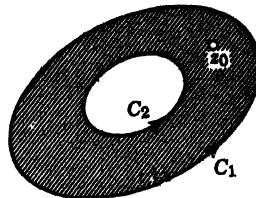


Fig. 327. Formula (3).

Since ϵ can be made arbitrarily small, it follows that the last integral on the right side of (2) has the value zero, and the theorem is proved.

Example 1. Integrate

$$\frac{z^2 + 1}{z^2 - 1}$$

along a circle of radius 1 with center at

- (a) $z = 1$ (b) $z = 1/2$ (c) $z = -1$ (d) $z = i$,

the integration being taken in the counterclockwise sense.

(a) The integral under consideration may be written as

$$\int_C \frac{z^2 + 1}{z^2 - 1} dz = \int_C \frac{z^2 + 1}{z + 1} \frac{dz}{z - 1}.$$

The expression on the right is of the form (1) where $z_0 = 1$ and

$$f(z) = \frac{z^2 + 1}{z + 1}.$$

The point $z_0 = 1$ lies inside the circle C under consideration, and $f(z)$ is analytic inside and on C . (The point $z = -1$, where $f(z)$ is not analytic, lies outside C .) Hence, by Cauchy's integral formula,

$$\int_C \frac{z^2 + 1}{z^2 - 1} dz = \int_C \frac{z^2 + 1}{z + 1} \frac{dz}{z - 1} = 2\pi i \left[\frac{z^2 + 1}{z + 1} \right]_{z=1} = 2\pi i.$$

(b) We obtain the same result as before because the given function is analytic except at the points $z = 1$ and $z = -1$, and we may obtain the circle (b) from that in the case (a) by a continuous deformation (even a translation) without passing over a point where the given function is not analytic.

(c) We may write

$$\int_C \frac{z^2 + 1}{z^2 - 1} dz = \int_C \frac{z^2 + 1}{z - 1} \frac{dz}{z + 1}.$$

The integral on the right is of the form (1) where $z_0 = -1$ and

$$f(z) = \frac{z^2 + 1}{z - 1}.$$

The point $z_0 = -1$ lies inside the circle C now under consideration, and $f(z)$ is analytic inside and on C . Therefore, by (1),

$$\int_C \frac{z^2 + 1}{z^2 - 1} dz = \int_C \frac{z^2 + 1}{z - 1} \frac{dz}{z + 1} = 2\pi i \left[\frac{z^2 + 1}{z - 1} \right]_{z=-1} = -2\pi i.$$

(d) The given function is analytic everywhere inside the circle now under consideration and on that circle. Hence, by Cauchy's integral theorem, the integral has the value zero.

In the case of a *multiply connected domain* we may proceed in a manner similar to that in Sec. 12.3. For instance, if $f(z)$ is analytic (and single-valued) on C_1 and C_2 and in the ring-shaped domain bounded by C_1 and C_2 (Fig. 327) and z_0 is any point in that domain, then

$$(3) \quad f(z_0) = \frac{1}{2\pi i} \int_{C_1} \frac{f(z)}{z - z_0} dz - \frac{1}{2\pi i} \int_{C_2} \frac{f(z)}{z - z_0} dz,$$

where both integrals are taken in the counterclockwise sense.

Problems

Integrate $z^2/(z^2 + 1)$ in the counterclockwise sense around the circle

$$\bullet \quad 1. |z + i| = 1 \quad 2. |z - i| = 1/2 \quad \bullet \quad 3. |z| = 2 \quad 4. |z| = 1/2$$

Integrate $z^2/(z^4 - 1)$ in the counterclockwise sense around the circle

$$5. |z - 1| = 1 \quad 6. |z + i| = 1 \quad 7. |z - i| = 1/2 \quad 8. |z| = 2$$

Integrate the following functions in the counterclockwise sense around the unit circle.

9. $1/z$	10. $1/(z^2 + 4)$	11. $1/(4z + i)$
12. e^z/z	13. $e^{2z}/(z + 2i)$	14. $e^{z^2}/(2z - i)$
15. $(\cos z)/z$	16. $(\sin z)/z$	17. $(e^z - 1)/z$
18. $(\sinh z)/z$	19. $(\cosh 3z)/z$	20. $(z^3 + 1)/(z^2 - 3iz)$

12.6 The Derivatives of an Analytic Function

From the assumption that a *real* function of a real variable is once differentiable nothing follows about the existence of derivatives of higher order. We shall now see that from the assumption that a *complex* function has a first derivative in a domain D there follows the existence of derivatives of all orders in D . This means that in this respect complex analytic functions behave much simpler than real functions which are once differentiable.

Theorem 1. *If $f(z)$ is analytic in a domain D , then it has derivatives of all orders in D which are then also analytic functions in D . The values of these derivatives at a point z_0 in D are given by the formulas*

$$(1') \quad f'(z_0) = \frac{1}{2\pi i} \int_C \frac{f(z)}{(z - z_0)^2} dz,$$

$$(1'') \quad f''(z_0) = \frac{2!}{2\pi i} \int_C \frac{f(z)}{(z - z_0)^3} dz,$$

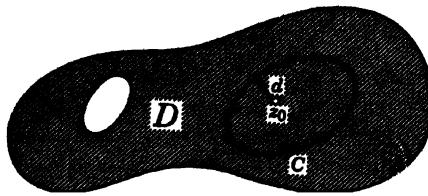


Fig. 328. Theorem 1.

and in general

$$(1) \quad f^{(n)}(z_0) = \frac{n!}{2\pi i} \int_C \frac{f(z)}{(z - z_0)^{n+1}} dz \quad (n = 1, 2, \dots);$$

here C is any simple closed path in D which encloses z_0 and whose full interior belongs to D ; the curve C is traversed in the counterclockwise sense (Fig. 328).

Remark. For memorizing (1), it is useful to observe that these formulas are obtained formally by differentiating the Cauchy formula (1), Sec. 12.5, under the integral sign with respect to z_0 .

Proof of Theorem 1. We prove (1'). By definition,

$$f'(z_0) = \lim_{\Delta z \rightarrow 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z}$$

From this and Cauchy's integral formula (1), Sec. 12.5, it follows that

$$(2) \quad f'(z_0) = \lim_{\Delta z \rightarrow 0} \frac{1}{2\pi i \Delta z} \left[\int_C \frac{f(z)}{z - (z_0 + \Delta z)} dz - \int_C \frac{f(z)}{z - z_0} dz \right]$$

Straightforward calculation shows that

$$\frac{1}{\Delta z} \left[\frac{1}{z - (z_0 + \Delta z)} - \frac{1}{z - z_0} \right] = \frac{1}{(z - z_0)^2} + \frac{\Delta z}{(z - z_0 - \Delta z)(z - z_0)^2}.$$

Hence, we may write (2) in the form

$$f'(z_0) = \frac{1}{2\pi i} \int_C \frac{f(z)}{(z - z_0)^2} dz + \lim_{\Delta z \rightarrow 0} \frac{\Delta z}{2\pi i} \int_C \frac{f(z)}{(z - z_0 - \Delta z)(z - z_0)^2} dz.$$

Consequently, (1') will be established if we show that the last expression on the right has the value zero. On C , $f(z)$ is continuous and, therefore, bounded in absolute value, say, $|f(z)| < M$. Let d be the distance of the point (or points) of C which are closest to z_0 . Then for all z on C ,

$$|z - z_0| \geq d \quad \text{and} \quad \frac{1}{|z - z_0|} \leq \frac{1}{d}.$$

Furthermore, if $|\Delta z| \leq d/2$, then for all these z ,

$$|z - z_0 - \Delta z| \geq \frac{d}{2} \quad \text{and} \quad \frac{1}{|z - z_0 - \Delta z|}$$

Denoting the length of C by L and using (4) in Sec. 12.2, we thus obtain

$$\frac{\Delta z}{2\pi i} \int_C \frac{f(z)}{(z - z_0 - \Delta z)(z - z_0)^n} dz \leq \frac{|\Delta z|}{2\pi} \frac{M}{L} \quad \left(|\Delta z| \leq \frac{d}{2} \right).$$

If Δz approaches zero, the expression on the right approaches zero. This proves (1'). Formula (1'') can be proved by a similar argument, using (1'), and the general formula (1) follows by induction.

Using Theorem 1, let us prove the converse of Cauchy's theorem.

Morera's⁵ theorem. *If $f(z)$ is continuous in a simply connected domain D and if*

$$(3) \quad \int_C f(z) dz = 0$$

for every closed path in D , then $f(z)$ is analytic in D .

Proof. In Sec. 12.4 it was shown that if $f(z)$ is analytic in D , then

$$F(z) = \int_{z_0}^z f(z^*) dz^*$$

is analytic in D and $F'(z) = f(z)$. In the proof we used only the continuity of $f(z)$ and the property that its integral around every closed path in D is zero; from these assumptions we concluded that $F(z)$ is analytic. By Theorem 1 the derivative of $F(z)$ is analytic, that is, $f(z)$ is analytic in D , and Morera's theorem is proved.

We shall now derive an important inequality. In (1), let C be a circle of radius r with center at z_0 , and let M be the maximum of $|f(z)|$ on C . Then, by applying (4), Sec. 12.2, to (1) we find

$$|f^{(n)}(z_0)| = \frac{n!}{2\pi} \left| \int_C \frac{f(z) dz}{(z - z_0)^{n+1}} \right| \leq \frac{n!}{2\pi} M \frac{1}{r^{n+1}} 2\pi r.$$

This yields Cauchy's inequality

$$(4) \quad |f^{(n)}(z_0)| \leq \frac{n! M}{r^n}.$$

Using (4), let us prove the following interesting and basic result.

Liouville's theorem. *If $f(z)$ is analytic and bounded in absolute value for all (finite) z , then $f(z)$ is a constant.*

Proof. By assumption, $|f(z)|$ is bounded, say, $|f(z)| < K$ for all z . Using (4), we see that $|f'(z_0)| < K/r$. Since this is true for every r , we can take r as large as we please and conclude that $f'(z_0) = 0$. Since z_0 is arbitrary, $f'(z) = 0$ for all finite z , and $f(z)$ is a constant (cf. Prob. 17 at the end of Sec. 10.3). This completes the proof.

⁵ GIACINTO MORERA (1856–1909), Italian mathematician. •

Problems

Using Theorem 1, integrate the following functions in the counterclockwise sense around the unit circle (n in Probs 8, 12, 13 is a positive integer)

- | | | |
|---------------------------|---------------------|--------------------------|
| 1. $z^2/(2z - 1)^2$ | 2. $z^2/(2z - 1)^4$ | 3. $z^3/(2z + i)^3$ |
| 4. $(z^3 + 1)/(3z + 1)^3$ | 5. $z/(4z + i)^2$ | 6. e^z/z^2 |
| 7. e^z/z^3 | 8. e^z/z^n | 9. $ze^z/(4z + \pi i)^2$ |
| 10. $z^{-2}\cos z$ | 11. $z^{-2}\sin z$ | 12. $z^{-2n}\cos z$ |
| 13. $z^{-2n}\cos z$ | 14. e^{z^2}/z^3 | 15. $z^{-2e^{-z}}\sin z$ |

16. Verify (4) for $f(z) = z^n$
17. If $f(z)$ is not a constant and is analytic for all (finite) z , and R and M are any positive real numbers (no matter how large), show that there exist values of z for which $|z| > R$ and $|f(z)| > M$. Hint: use Liouville's theorem.
18. If $f(z)$ is a polynomial of degree $n > 0$ and M an arbitrary positive real number (no matter how large) show that there exists a positive real number R such that $|f(z)| > M$ for all $|z| > R$.
19. Show that $f(z) = e^z$ has the property characterized in Prob 17, but does not have that characterized in Prob 18.
20. Prove the **Fundamental theorem of algebra**: If $f(z)$ is a polynomial in z , not a constant, then $f(z) = 0$ for at least one value of z . Hint: assume $f(z) \neq 0$ for all z and apply the result of Prob 17 to $g = 1/f$.

Chapter 13

Sequences and Series

Infinite sequences and series play a basic role in complex analysis and its applications. We shall define sequences, series, and related notions, consider tests for convergence and divergence, and show that power series represent analytic functions. The fact that, conversely, analytic functions can always be represented by power series will be considered in the next chapter. The notions appearing in the present chapter will be similar to those in real analysis, but the chapter is self-contained and independent.

Prerequisites for this chapter: Secs. 10.1–10.3.

References: Appendix 1, Part G.

Answers to problems: Appendix 2.

13.1 Sequences

Infinite sequences and series play a basic role in complex analysis and its applications. We shall see that most of the definitions and theorems relating to complex sequences and series are quite similar to those in the case of real sequences and series.

If to each positive integer n there is assigned a number z_n , then these numbers

$$z_1, z_2, \dots, z_n, \dots$$

are said to form an *infinite sequence* or, briefly, a **sequence**, and the numbers z_n are called the **terms** of the sequence.

A sequence whose terms are real is called a *real sequence*.

At times it is convenient to number the terms of a sequence starting with 0, with 2, or with some other integer.

A sequence z_1, z_2, \dots is said to *converge to the number c*, or to *be convergent with the limit c*, if to each positive number ϵ (no matter how small but not zero) we can find a number N such that

$$(1) \quad |z_n - c| < \epsilon \quad \text{for each } n > N,$$

and then we write

$$\lim_{n \rightarrow \infty} z_n = c.$$

A sequence which is not convergent is said to *be divergent* or to *diverge*.

The condition (1) has a simple geometrical interpretation. It means that each term z_n with $n > N$ lies in the open circular disk of radius ϵ with center at c (Fig. 329), and at most finitely many terms z_n do not lie in that disk, no matter how small the radius ϵ of the disk is chosen. Of course, N will depend on the choice of ϵ , in general.

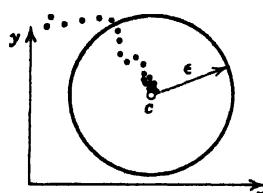


Fig. 329. Convergent complex sequence.

In the case of a *real* sequence, the condition (1) means geometrically that each term z_n with $n > N$ lies between $c - \epsilon$ and $c + \epsilon$ (Fig. 330), and at most finitely many terms of the sequence do not lie in that interval.

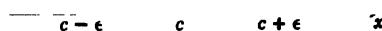


Fig. 330. Convergent real sequence.

Example 1. The sequence whose terms are $z_n = 1 + \frac{2}{n}$, ($n = 1, 2, \dots$), that is,

$$3, 2, \frac{5}{3}, \frac{8}{4}, \frac{7}{5}, \dots$$

is convergent with the limit $c = 1$. In fact, in (1),

$$z_n - c = 1 + \frac{2}{n} - 1 = \frac{2}{n} \quad \text{and} \quad \frac{2}{n} < \epsilon \quad \text{when} \quad \frac{n}{2} > \frac{1}{\epsilon} \quad \text{or} \quad n > \frac{2}{\epsilon}.$$

For example, choosing $\epsilon = \frac{1}{100}$ we have $\frac{2}{n} < \frac{1}{100}$ when $n > 200$.

Example 2. The sequences $1, 2, 3, \dots$ and $\frac{1}{4}, \frac{2}{4}, \frac{1}{5}, \frac{4}{5}, \frac{1}{6}, \frac{5}{6}, \dots$ are divergent.

Example 3. The sequence whose terms are $z_n = \frac{2n-1}{n} + i \frac{n+2}{n}$ is

$$1 + 3i, \quad \frac{3}{2} + 2i, \quad \frac{5}{3} + \frac{5}{3}i, \quad \frac{7}{4} + \frac{3}{2}i, \dots$$

It converges to $c = 2 + i$. In fact, in (1),

$$|z_n - c| = \left| \frac{2n-1}{n} + i \frac{n+2}{n} - (2+i) \right| = \left| -\frac{1}{n} + \frac{2i}{n} \right| = \frac{\sqrt{5}}{n}$$

and $\frac{\sqrt{5}}{n} < \epsilon$ when $\frac{n}{\sqrt{5}} > \frac{1}{\epsilon}$ or $n > \frac{\sqrt{5}}{\epsilon} = \frac{2.236 \dots}{\epsilon}$.

For example, choosing $\epsilon = \frac{1}{225}$ we thus obtain

$$|z_n - c| < \epsilon \quad \text{when} \quad n = 224, 225, \dots$$

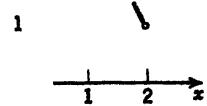


Fig. 331. Sequence in Example 3.

A complex sequence z_1, z_2, \dots being given, we may consider the sequence of the real parts and the sequence of the imaginary parts. Setting

$$z_n = x_n + iy_n,$$

these sequences are

$$x_1, x_2, x_3, \dots \quad \text{and} \quad y_1, y_2, y_3, \dots$$

For instance, in Ex. 3 the two sequences are

$$1, \frac{3}{2}, \frac{5}{3}, \frac{7}{4}, \dots \quad \text{and} \quad 3, 2, \frac{5}{2}, \frac{3}{2}, \dots,$$

and we observe that the sequence of the real parts converges to 2, the real part of c , while the sequence of the imaginary parts converges to 1, the imaginary part of c . This illustrates

Theorem 1. A sequence $z_1, z_2, \dots, z_n, \dots$ of complex numbers $z_n = x_n + iy_n$ ($n = 1, 2, \dots$) converges to the limit $c = a + ib$ if and only if the sequence of the real parts x_1, x_2, \dots converges to a and the sequence of the imaginary parts y_1, y_2, \dots converges to b .

Proof. If $|z_n - c| < \epsilon$, then $z_n = x_n + iy_n$ is within the circle of radius ϵ about $c = a + ib$ so that necessarily (Fig. 332a)

$$|x_n - a| < \epsilon, \quad |y_n - b| < \epsilon.$$

Thus convergence of z_n to c as $n \rightarrow \infty$ implies convergence of x_n to a and y_n to b .

Conversely, if $x_n \rightarrow a$ and $y_n \rightarrow b$ as $n \rightarrow \infty$, then for given $\epsilon > 0$ we can choose N so large that, for every $n > N$,

$$|x_n - a| < \frac{\epsilon}{2}, \quad |y_n - b| < \frac{\epsilon}{2}.$$

These two inequalities imply that $z_n = x_n + iy_n$ lies in a square with center c and side ϵ . Hence, z_n must lie within a circle of radius ϵ about c (Fig. 332b). This completes the proof.

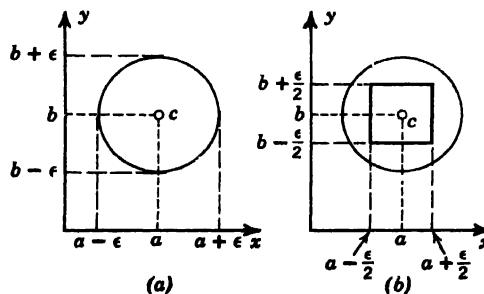


Fig. 332. Proof of Theorem 1.

This theorem shows that, by studying the real and imaginary parts, the convergence of complex sequences can be referred back to that of real sequences.

A point a is called a **limit point** of a sequence z_1, z_2, \dots if, given an $\epsilon > 0$ (no matter how small),

$$(2) \quad |z_n - a| < \epsilon \quad \text{for infinitely many } n.$$

Geometrically speaking, this means that infinitely many terms of the sequence lie within the circle of radius ϵ about a , no matter how small the radius ϵ is chosen.

Note that if (2) holds, there may still be infinitely many terms which do not lie within that circle, and the sequence may not be convergent. In fact, if a sequence has more than one limit point, it diverges. The following example may illustrate the situation.

Example 4.

Sequence	Limit points at:	Convergent or divergent
1, 2, 3, ...	(none)	divergent
$\frac{1}{2}, \frac{8}{3}, \frac{3}{4}, \frac{4}{5}, \dots$	1	convergent
$\frac{1}{2}, 2, \frac{1}{3}, 3, \frac{1}{4}, 4, \dots$	0	divergent
$\frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{4}{5}, \frac{1}{6}, \frac{5}{6}, \dots$	0 and 1	divergent

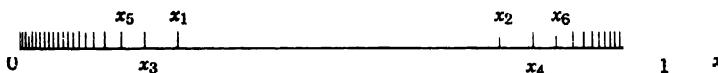


Fig. 333. Last sequence in Example 4.

A number which appears infinitely often in a sequence is to be regarded as a limit point; this is a matter of convenience and convention.

A sequence z_1, z_2, \dots is said to be **bounded**, if there is a positive number K such that all the terms of the sequence lie in a disk of radius K about the origin, that is,

$$|z_n| < K \quad \text{for all } n.$$

For example, the second and the last sequences in Ex. 4 are bounded while the first and third are not. We observe that the two bounded sequences have limit points. This illustrates the following important theorem.

Theorem 2 (Bolzano¹ and Weierstrass²). *A bounded infinite sequence has at least one limit point.*

Proof. It is obvious that both conditions are necessary: a finite sequence cannot have a limit point, and the sequence 1, 2, 3, ..., though infinite, has no limit point because it is not bounded. To prove the theorem, consider a bounded infinite sequence z_1, z_2, \dots and let K be such that $|z_n| < K$ for all n . If only finitely many values of the z_n are different, then, since the sequence is infinite, some number z must occur infinitely many times in the sequence, and, by definition, this number is a limit point of the sequence.

¹ BERNHARD BOLZANO (1781–1848), German mathematician, a pioneer in the study of point sets.

² Cf. footnote 3 in Sec. 10.3.

We may now turn to the case when the sequence contains infinitely many different terms. We draw the large square Q_0 in Fig. 334 which contains all z_n . We subdivide Q_0 into four congruent squares. Clearly, at least one of these squares (each taken with its complete boundary) must contain infinitely many terms of the sequence. The square of this type with the lowest number (1, 2, 3, or 4) will be denoted by Q_1 . This is the first step. In the next step we subdivide Q_1 into four congruent squares and select a square Q_2 according to the same rule, and so on. This yields an infinite sequence of squares $Q_0, Q_1, Q_2, \dots, Q_n, \dots$ with the property that the side of Q_n approaches zero as n approaches infinity, and Q_m contains all Q_n with $n > m$. It is not difficult to see that the number which belongs to all these squares,³ call it $z = a$, is a limit point of the sequence. In fact, given an $\epsilon > 0$, we can choose an N so large that the side of the square Q_N is less than ϵ and, since Q_N contains infinitely many z_n , we have $|z_n - a| < \epsilon$ for infinitely many n . This completes the proof.

By the use of this theorem we shall now establish a general criterion for convergence which will be fundamental in our further consideration. The point is that this criterion can be applied without knowing in advance the value of the limit. This is a great advantage, compared with the direct application of the definition of a convergent sequence.

Theorem 3 (Cauchy's convergence principle). *A sequence z_1, z_2, \dots is convergent if, and only if, for every positive number ϵ we can find a number N (which may depend on ϵ) such that*

$$(3) \quad |z_m - z_n| < \epsilon \quad \text{when } m > N, n > N;$$

(that is, any two terms z_m, z_n with $m > N, n > N$ must have a distance of less than ϵ from each other).

Proof. (a) Let the sequence z_1, z_2, \dots be convergent with the limit c . Then, for a given $\epsilon > 0$, we can find an N such that

$$|c| < \frac{\epsilon}{2} \quad \text{for every } n > N.$$

Hence, when $m > N, n > N$, then by the triangle inequality

$$|z_m - z_n| = |(z_m - c) - (z_n - c)| \leq |z_m - c| + |z_n - c| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

which shows that, if the sequence converges, then (3) holds.

³ The fact that such a unique number $z = a$ exists seems to be obvious, but it actually follows from an axiom of the real number system, the so-called *Cantor-Dedekind axiom*. Cf. the next footnote.

(b) Conversely, consider a sequence z_1, z_2, \dots for which (3) holds. We first show that the sequence is bounded. In fact, choose a fixed ϵ and a fixed $n = n_0 > N$ in (3). Then (3) implies that every z_m with $m > N$ lies in a disk of radius ϵ about z_{n_0} , and only finitely many terms of the sequence do not lie in that disk. It is clear that we can find a circle about the origin so large that the disk and those finitely many z_n lie within the circle. This shows that the sequence is bounded, and from the Bolzano-Weierstrass theorem we conclude that it has at least one limit point, call it L .

We show that the sequence is convergent with the limit L . The definition of a limit point implies that, an $\epsilon > 0$ being given, $|z_n - L| < \epsilon/2$ for infinitely many n . Since (3) holds for any $\epsilon > 0$, when an $\epsilon > 0$ is given we can find an N^* such that $|z_m - z_n| < \epsilon/2$ for $m > N^*, n > N^*$. Choose a fixed $n > N^*$ such that $|z_n - L| < \epsilon/2$, and let m be any integer greater than N^* . Then, by the triangle inequality,

$$|z_m - L| = |(z_m - z_n) + (z_n - L)| \leq |z_m - z_n| + |z_n - L| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

that is, $|z_m - L| < \epsilon$ for all $m > N^*$; by definition, this means that the sequence is convergent with the limit L . This completes the proof.

This theorem will play an important role in our considerations of infinite series.

The following property of a convergent sequence is almost obvious.

Theorem 4. *A convergent sequence is bounded. Consequently, if a sequence is not bounded, it diverges.*

Proof. Let the sequence z_1, z_2, \dots be convergent with the limit c . Then we may choose an $\epsilon > 0$ and find a corresponding N such that every z_n with $n > N$ lies in the disk of radius ϵ with center c , and at most finitely many terms may not lie in that disk. It is clear that we can find a circle of radius K about the origin so large that the disk as well as those finitely many terms lie within that circle. This proves that the sequence is bounded.

Note well that the condition of Theorem 4 is only necessary for convergence but not sufficient. This is illustrated by the last sequence in Ex. 4, which is bounded but has two limit points and, therefore, is divergent. The first and third sequences in Ex. 4 are not bounded, and from Theorem 4 it follows that they are divergent.

In the case of a *real sequence* we may find a simple condition sufficient for convergence. For this purpose we need the following concepts.

A real sequence $x_1, x_2, \dots, x_n, \dots$ is said to be **monotone increasing** if

$$x_1 \leq x_2 \leq x_3 \leq \dots$$

Similarly, the sequence is said to be **monotone decreasing** if

$$x_1 \geq x_2 \geq x_3 \geq \dots$$

A sequence which is either monotone increasing or monotone decreasing is called a *monotone sequence*.

For example, the first two sequences in Ex. 4 are monotone, while the last two are not monotone. We observe that the second sequence in Ex. 4 is bounded and monotone, and we shall now prove that these two properties together are sufficient for convergence.

Theorem 5. *If a real sequence is bounded and monotone, it converges.*

Proof. Let x_1, x_2, \dots be a bounded monotone increasing sequence. Then its terms are smaller than some number B and, since $x_1 \leq x_n$ for all n , they lie in the interval $x_1 \leq x_n \leq B$, which will be denoted by I_0 . We bisect I_0 , that is, we subdivide it into two parts of equal length. If the right half (together with its end points) contains terms of the sequence, we denote it by I_1 . If it does not contain terms of the sequence, then the left half of I_0 (together with its end points) is called I_1 . This is the first step.

In the second step we bisect I_1 , select one half by the same rule, and call it I_2 , and so on (Fig. 335).

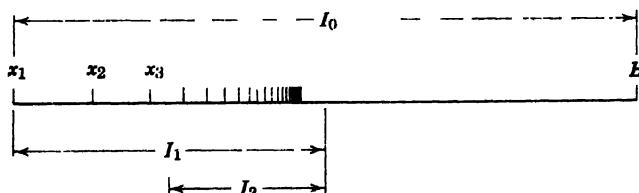


Fig. 335. Proof of Theorem 5.

In this way we obtain shorter and shorter intervals I_0, I_1, I_2, \dots with the following properties. Each I_m contains all I_n for $n > m$. No term of the sequence lies to the right of I_m , and, since the sequence is monotone increasing, all x_n with n greater than some number N lie in I_m ; of course, N will depend on m , in general. The lengths of the I_m approach zero as m approaches infinity. Hence, there is precisely one number, call it L , which lies in all those intervals,⁴ and we may now easily prove that the sequence is convergent with the limit L .

In fact, given an $\epsilon > 0$, we choose an m such that the length of I_m is less than ϵ . Then L and all the x_n with $n > N(m)$ lie in I_m , and, therefore, $|x_n - L| < \epsilon$ for all those n . This completes the proof for an increasing

⁴ This statement seems to be obvious, but actually it is not; it may be regarded as an axiom of the real number system in the following form. Let J_1, J_2, \dots be closed intervals such that each J_m contains all J_n with $n > m$, and the lengths of the J_m approach zero as m approaches infinity. Then there is precisely one real number which is contained in all those intervals. This is the so-called Cantor-Dedekind Axiom, named after the German mathematicians, GEORG CANTOR (1845–1918), the creator of set theory, and RICHARD DEDEKIND (1831–1916). For further details see Ref. [A9] in Appendix 1. (An interval I is said to be closed if its two end points are regarded as points belonging to I . It is said to be open if the end points are not regarded as points of I .)

sequence For a decreasing sequence the proof is the same, except for a suitable interchange of "left" and "right" in the construction of those intervals

Problems

Write the first few terms of the sequence z_1, z_2, \dots where

1. $z_n = n/(n+3)$
2. $z_n = 2n/(n^2 + 1)$
3. $z_n = i^n/n^3$
4. $z_n = (-1)^n/n^2$
5. $z_n = (-1)^{n+1}/(n+1)$
6. $z_n = (2i)^n/3^{n-1}$
7. $z_1 = 1, z_2 = 2, z_3 = 3, z_n = z_{n-3} - z_{n-2} + z_{n-1} (n = 4, 5, \dots)$
8. $z_1 = 1, z_2 = 1/2, z_n = z_{n-2}z_{n-1} (n = 3, 4, \dots)$
9. $z_1 = 1/3, z_2 = 1/4, z_n = z_{n-2}/z_{n-1} (n = 3, 4, \dots)$
10. $z_1 = 1, z_2 = i, z_n = z_{n-2}z_{n-1} (n = 3, 4, \dots)$

Are the following sequences bounded or not, convergent or divergent? Determine their limit points

11. $z_n = i^n$
12. $z_n = i^n/n$
13. $z_n = in/(n+1)$
14. $z_n = n^2/(n+1)$
15. $z_n = (-1)^n/n^3$
16. $z_n = e^{in\pi/4}$
17. $z_n = \pi i + \sin n\pi$
18. $z_n = (\cos n\pi)/\sqrt{n}$
19. $z_n = i^n n^2$

Are the following real sequences bounded or not, convergent or not, monotone or not? Determine their limit points

20. $1, \frac{1}{3}, \frac{1}{9}, \frac{1}{27}, \dots$
21. $2, -\frac{1}{2}, 3, -\frac{1}{3}, 4, -\frac{1}{4}, \dots$
22. $2, 2^2, 2^3, 2^4, \dots$
23. $1, \sqrt{2}, \sqrt[3]{3}, \sqrt[4]{4}, \dots$
24. $\frac{7}{6}, \frac{11}{8}, \frac{8}{7}, \frac{13}{9}, \frac{9}{8}, \frac{15}{14}, \dots$
25. $\ln 1, \ln 2, \ln 3, \dots$
26. $4/1!, 4^2/2!, 4^3/3!, \dots$
27. a, a^2, a^3, \dots
28. $c, 2c^2, 3c^3, \dots (|c| < 1)$
29. $c, 2^2c^2, 3^2c^3, 4^2c^4, \dots (|c| < 1)$

30. If the sequence z_1, z_2, \dots is convergent with the limit 0, and the sequence b_1, b_2, \dots is such that $|b_n| \leq K|z_n|$ for some fixed $K > 0$ and all n , show that the sequence b_1, b_2, \dots is convergent with the limit 0

13.2 Series

Let $w_1, w_2, \dots, w_n, \dots$ be a sequence of numbers, complex, or real. Then we may consider the infinite series

$$(1) \quad \sum_{m=1}^{\infty} w_m = w_1 + w_2 + w_3 + \dots$$

The w_m are called the terms of the series. The sum of the first n terms is

$$(2) \quad s_n = w_1 + w_2 + \dots + w_n$$

This expression is called the n th partial sum of the series (1). Clearly if we omit the terms of s_n from (1), the remaining expression is

$$(3) \quad R_n = w_{n+1} + w_{n+2} + w_{n+3} + \dots$$

This is called the remainder of the series (1) after the n -th term.

In this way we have now associated with the series (1) the sequence of its partial sums s_1, s_2, s_3, \dots . If this sequence is convergent, say,

$$\lim_{n \rightarrow \infty} s_n = s,$$

then the series (1) is said to **converge** or to be *convergent*, the number s is called its **value** or **sum**, and we write

$$s = \sum_{m=1}^{\infty} w_m = w_1 + w_2 + \dots$$

If the sequence of the partial sums diverges, the series (1) is said to **diverge** or to be *divergent*.

If the series (1) converges and has the value s , then

$$(4) \quad s = s_n + R_n \quad \text{or} \quad R_n = s - s_n.$$

Example 1. The series

$$\sum_{m=1}^{\infty} \frac{1}{2^m} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots$$

converges, because

$$s_n = \frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2^n} = 1 - \frac{1}{2^n} \quad \text{and} \quad \lim_{n \rightarrow \infty} s_n = 1;$$

the series has the value 1.

Example 2. The series

$$\sum_{m=1}^{\infty} m = 1 + 2 + 3 + \dots$$

diverges.

Example 3. The series

$$\sum_{m=0}^{\infty} (-1)^m = 1 - 1 + 1 - + \dots$$

diverges, because

$$s_0 = 1, \quad s_1 = 1 - 1 = 0, \quad s_2 = 1 - 1 + 1 = 1, \quad \text{etc.},$$

and the sequence $1, 0, 1, 0, \dots$ is divergent.

Example 4. The harmonic series

$$\sum_{m=1}^{\infty} \frac{1}{m} = 1 + \frac{1}{2} + \frac{1}{3} + \dots$$

diverges. In fact,

$$s_n = 1 + \frac{1}{2} + \dots + \frac{1}{n},$$

and s_n equals the sum of the areas of the n rectangles in Fig. 336 on the next page. This area is greater than the area A_n under the corresponding portion of the curve $y = 1/x$, and

$$A_n = \int_1^{n+1} \frac{dx}{x} = \ln(n+1) \rightarrow \infty \quad \text{as } n \rightarrow \infty.$$

Since $s_n > A_n$, it follows that $s_n \rightarrow \infty$ as $n \rightarrow \infty$, which means divergence.

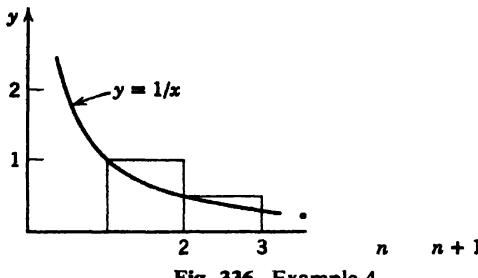


Fig. 336. Example 4.

Theorem 1. If the series $w_1 + w_2 + \cdots$ converges, then

$$(5) \quad \lim_{m \rightarrow \infty} w_m = 0.$$

Consequently, a series which does not satisfy (5) diverges.

Proof. Let $w_1 + w_2 + \cdots$ be convergent with the sum s . Then

$$w_{n+1} = s_{n+1} - s_n,$$

and

$$\lim_{n \rightarrow \infty} (s_{n+1} - s_n) = \lim_{n \rightarrow \infty} s_{n+1} - \lim_{n \rightarrow \infty} s_n = s - s = 0.$$

Note well that the condition (5) is only necessary for convergence, but not sufficient. In fact, the harmonic series satisfies (5), but diverges. The series in Exs. 2 and 3 do not satisfy (5) and, therefore, are divergent.

If a given series $w_1 + w_2 + \cdots$ converges and has the value s , then

$$s = s_n + R_n \quad \text{or} \quad R_n = s - s_n \quad [\text{cf. (4)}]$$

and from the definition of convergence of a sequence it follows that $|R_n|$ can be made as small as we please, by taking n large enough. In many cases it will be impossible to find the sum s of a convergent series. Then for computational purposes we must use a partial sum s_n as an approximation of s , and by estimating the remainder R_n we may obtain information about the degree of accuracy of the approximation.

In the case of a *real series* whose terms have alternating sign and decrease in absolute value, the following useful theorem may be applied for that purpose.

Theorem 2 (Leibniz test for real series). Let u_1, u_2, \dots be real and

$$(6) \quad (a) \quad u_1 \geq u_2 \geq u_3 \geq \cdots, \quad (b) \quad \lim_{m \rightarrow \infty} u_m = 0.$$

Then the series

$$u_1 - u_2 + u_3 - u_4 + - \cdots$$

converges, and for the remainder R_n after the n -th term we have the estimate

$$(7) \quad |R_n| \leq u_{n+1}.$$

$$\begin{array}{c} \text{--- } -u_2 \\ \text{--- } u_3 \text{ ---} \end{array}$$

s_2 s_4 s_3 s_1

Fig. 337. Proof of Theorem 2 (Leibniz test).

Proof. Let s_n be the n th partial sum of the series. Then, because of (6a),

$$\begin{aligned} s_1 &= u_1, & s_2 &= u_1 - u_2 \leq s_1, \\ s_3 &= s_2 + u_3 \geq s_2, & s_4 &= s_1 - (u_2 - u_3) \leq s_1, \end{aligned}$$

so that $s_2 \leq s_3 \leq s_1$. Proceeding in this fashion, we conclude that (Fig. 337)

$$(8) \quad s_1 \geq s_3 \geq s_5 \geq \cdots \geq s_6 \geq s_4 \geq s_2$$

which shows that the odd partial sums form a bounded monotone sequence, and so do the even partial sums. Hence, by Theorem 5, Sec. 13.1, both sequences converge, say,

$$\lim_{n \rightarrow \infty} s_{2n+1} = s, \quad \lim_{n \rightarrow \infty} s_{2n} = s^*.$$

Now, since $s_{2n+1} - s_{2n} = u_{2n+1}$, we obtain from (6b)

$$s - s^* = \lim_{n \rightarrow \infty} s_{2n+1} - \lim_{n \rightarrow \infty} s_{2n} = \lim_{n \rightarrow \infty} (s_{2n+1} - s_{2n}) = \lim_{n \rightarrow \infty} u_{2n+1} = 0.$$

Hence $s^* = s$, and the series converges with the sum s .

We prove (7). Since $s_n \rightarrow s$, it follows from (8) that

$$s_{2n+1} \geq s \geq s_{2n} \quad \text{and also} \quad s_{2n-1} \geq s \geq s_{2n}.$$

By subtracting s_{2n} and s_{2n-1} , respectively, we obtain

$$s_{2n+1} - s_{2n} \geq s - s_{2n} \geq 0, \quad 0 \geq s - s_{2n-1} \geq s_{2n} - s_{2n-1}.$$

In these inequalities, the first expression is equal to u_{2n+1} , the last is equal to $-u_{2n}$, and the expressions between the inequality signs are the remainders R_{2n} and R_{2n-1} . Thus the inequalities may be written

$$u_{2n+1} \geq R_{2n} \geq 0, \quad 0 \geq R_{2n-1} \geq -u_{2n}$$

and imply (7). This completes the proof of Theorem 2.

Example 5. The series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + - \cdots$$

is convergent, as follows from Theorem 2.

From Theorem 1, Sec. 13.1, we immediately obtain the following relation between a complex series and the series of its real and imaginary parts.

Theorem 3. Let $w_m = u_m + iv_m$. Then the series

$$\sum_{m=1}^{\infty} w_m = w_1 + w_2 + w_3 + \cdots$$

converges and has the value $s = a + ib$ if, and only if, the series of the real parts and the series of the imaginary parts, that is

$$\sum_{m=1}^{\infty} u_m = u_1 + u_2 + \cdots \quad \text{and} \quad \sum_{m=1}^{\infty} v_m = v_1 + v_2 + \cdots,$$

converge and have the values a and b , respectively.

Furthermore, we may apply Theorem 3 in Sec. 13.1 to the sequence of the partial sums s_n of a given series $w_1 + w_2 + \cdots$. Then the inequality in that theorem becomes

$$|s_m - s_n| < \epsilon \quad (m > N, n > N)$$

or, if we set $m = n + p$,

$$|s_{n+p} - s_n| < \epsilon \quad (n > N, p = 1, 2, \dots).$$

Now by definition of a partial sum,

$$s_{n+p} - s_n = w_{n+1} + w_{n+2} + \cdots + w_{n+p}$$

and we immediately obtain the following basic result.

Theorem 4. A series $w_1 + w_2 + \cdots$ is convergent if, and only if, for any given $\epsilon > 0$ (no matter how small) we can find an N (which will depend on ϵ , in general) such that

$$|w_{n+1} + w_{n+2} + \cdots + w_{n+p}| < \epsilon \quad \text{for every } n > N \text{ and } p = 1, 2, \dots.$$

Theorem 3 establishes a relation between complex and real series. A far more important relation can be gained by the use of the following basic notion.

A series $w_1 + w_2 + \cdots$ is said to be **absolutely convergent** if the corresponding series

$$(9) \quad \sum_{m=1}^{\infty} |w_m| = |w_1| + |w_2| + \cdots$$

(whose terms are real and nonnegative) converges.

If the series $w_1 + w_2 + \cdots$ converges but (9) diverges, then the series is called, more precisely, *conditionally convergent*.

For instance, the series in Ex. 5 is conditionally convergent because the corresponding series (9) is the harmonic series (Ex. 4), which diverges. The series

$$\frac{1}{2} - \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots$$

is absolutely convergent, because the corresponding series (9) is convergent (cf. Ex. 1).

The following property of an absolutely convergent series is almost obvious.

Theorem 5. If a series $w_1 + w_2 + \cdots$ is absolutely convergent, it is convergent.

Proof. By the generalized triangle inequality (3) in Sec. 10.1 we have

$$(10) \quad |w_{n+1} + \cdots + w_{n+p}| \leq |w_{n+1}| + |w_{n+2}| + \cdots + |w_{n+p}|.$$

Since, by assumption the series $|w_1| + |w_2| + \cdots$ converges, it follows from Theorem 4 that the right-hand side of (10) becomes less than any given $\epsilon > 0$ for each $n > N$ (N sufficiently large) and $p = 1, 2, \dots$. Hence, the same is true for the left-hand side of (10), and by the same theorem, this proves convergence of the series $w_1 + w_2 + \cdots$.

13.3 Tests for Convergence and Divergence of Series

Of course, before we can use a series for computational or other purposes we must know whether it converges. In most cases that arise in engineering mathematics, this question may be answered by applying one of the various tests for convergence and divergence. These tests, therefore, are of great practical interest.

A simple test for divergence (Theorem 1 in the last section) and the so-called Leibniz test for real series were already considered. The following theorem is a source of various criteria for convergence.

Theorem 1 (Comparison test). If a series $w_1 + w_2 + \cdots$ is given and we can find a converging series $b_1 + b_2 + \cdots$ with non-negative real terms such that

$$(1) \quad |w_n| \leq b_n \quad \text{for } n = 1, 2, \dots,$$

then the given series is absolutely convergent.

Proof. Since the series $b_1 + b_2 + \cdots$ converges, it follows from Theorem 4 in the last section that for any given $\epsilon > 0$ we can find an N such that

$$b_{n+1} + \cdots + b_{n+p} < \epsilon \quad \text{for every } n > N \text{ and } p = 1, 2, \dots.$$

From this and (1) we conclude that

$$|w_{n+1}| + \cdots + |w_{n+p}| \leq b_{n+1} + \cdots + b_{n+p} < \epsilon$$

for those n and p . Hence, by Theorem 4 in the last section, the series $|w_1| + |w_2| + \cdots$ is convergent, and the given series converges absolutely.

For deriving two important tests from Theorem 1 let us prove

Theorem 2. The geometric series

$$\sum_{m=0}^{\infty} q^m = 1 + q + q^2 + \cdots$$

converges with the sum $1/(1 - q)$ when $|q| < 1$ and diverges when $|q| \geq 1$.

Proof. When $|q| \geq 1$, then $|q|^n \geq 1$ and divergence follows from Theorem 1 in the last section. Now let $|q| < 1$. The n th partial sum is

$$s_n = 1 + q + \cdots + q^n.$$

From this, $qs_n = q + \cdots + q^n + q^{n+1}$.

By subtraction,

$$s_n - qs_n = (1 - q)s_n = 1 - q^{n+1}$$

since the other terms cancel in pairs. Now $q - 1 \neq 0$ because $q \neq 1$, and we may solve for s_n , finding

$$(2) \quad s_n = \frac{1 - q^{n+1}}{1 - q} = \frac{1}{1 - q} - \frac{q^{n+1}}{1 - q}.$$

Since $|q| < 1$, the last term approaches zero as $n \rightarrow \infty$. Hence the series is convergent and has the value $1/(1 - q)$. This completes the proof.

From Theorems 1 and 2 we shall now derive two important tests, the ratio test and the root test.

Theorem 3 (Ratio test). Suppose that $w_n \neq 0$ for $n = 1, 2, \dots$ and the sequence

$$\left| \frac{w_{n+1}}{w_n} \right| \quad n = 1, 2,$$

converges with the limit L . Then the series

$$w_1 + w_2 + w_3 + \cdots$$

is

absolutely convergent, if $L < 1$,
divergent, if $L > 1$.

(If $L = 1$, the test fails).

Proof. By assumption

$$\lim_{n \rightarrow \infty} b_n = L \quad \text{where} \quad b_n = \left| \frac{w_{n+1}}{w_n} \right|.$$

Clearly, b_n and L are real. By the definition of limit, for given $\epsilon > 0$ we can find an N such that every b_n with $n > N$ lies between $L - \epsilon$ and $L + \epsilon$, that is,

$$(3) \quad (a) \quad b_n < L + \epsilon \quad (b) \quad b_n > L - \epsilon \quad (n > N).$$

We first consider the case $L < 1$. We set $L + \epsilon = q$ and choose $\epsilon = (1 - L)/2$. Then $\epsilon > 0$, and (3a) becomes

$$b_n < q = L + \frac{1 - L}{2} = \frac{1 + L}{2}.$$

Since $L < 1$, we have $q < 1$. Now we can write

$$(4) \quad \begin{aligned} |w_{N+1}| + |w_{N+2}| + |w_{N+3}| + \cdots \\ = |w_{N+1}| \left(1 + \left| \frac{w_{N+2}}{w_{N+1}} \right| + \left| \frac{w_{N+3}}{w_{N+2}} \right| \left| \frac{w_{N+2}}{w_{N+1}} \right| + \right. \\ \left. = |w_{N+1}| (1 + b_{N+1} + b_{N+2}b_{N+1} + b_{N+3}b_{N+2}b_{N+1} + \cdots). \right. \end{aligned}$$

Since $b_n < q < 1$, each term of this series is less than the corresponding term of the geometric series

$$|w_{N+1}| (1 + q + q^2 + q^3 + \cdots).$$

By Theorem 2 this series converges because $q < 1$. From Theorem 1 it follows that the series in (4) is convergent. From this we conclude that the series $|w_1| + |w_2| + \cdots$ is convergent. This implies that the series $w_1 + w_2 + \cdots$ is absolutely convergent.

We now consider the case $L > 1$. We choose $\epsilon = (L - 1)/2$. Clearly, $\epsilon > 0$, and (3b) becomes

$$b_n > L - \epsilon = \frac{1+L}{2} > 1 \quad (n > N),$$

that is,

$$b_n = \left| \frac{w_{n+1}}{w_n} \right| > 1 \quad \text{or} \quad |w_{n+1}| > |w_n| \quad (n > N).$$

The last inequality shows that the terms are increasing in absolute value. From this and Theorem 1, Sec. 13.2, it follows that the series is divergent.

For $L = 1$ the series may converge or diverge so that the test fails. This is illustrated by the harmonic series (Ex. 4 in Sec. 13.2)

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots$$

which diverges and for which

$$\frac{w_{n+1}}{w_n} = \frac{n}{n+1} \rightarrow L = 1 \quad \text{as } n \rightarrow \infty$$

and by the series

$$(5) \quad \sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{4} + \frac{1}{9} + \cdots$$

which converges and for which

$$\frac{w_{n+1}}{w_n} = \frac{n^2}{(n+1)^2} \rightarrow L = 1 \quad \text{as } n \rightarrow \infty.$$

Convergence of (5) may be shown as follows. The n th partial sum is

$$s_n = 1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{n^2}.$$

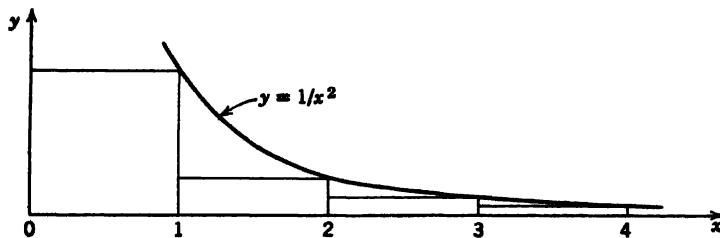


Fig. 338. Convergence of the series (5).

Clearly, $s_n > 0$ and (Fig. 338)

$$s_n \leq 1 + \int_1^n \frac{dx}{x^2} = 2 - \frac{1}{n},$$

which shows that the sequence of the partial sums is bounded. Since the terms of the series are positive, the sequence is monotone increasing, and convergence follows from Theorem 5 in Sec. 13.1.

The following test is more general than the ratio test, but its application is often more difficult.

Theorem 4 (Root test). Suppose that the sequence

$$\sqrt[n]{|w_n|} \quad n = 1, 2,$$

is convergent with the limit L . Then the series

$$w_1 + w_2 + w_3 + \cdots$$

is

absolutely convergent, if $L < 1$,
divergent, if $L > 1$.

(If $L = 1$, the test fails).

Proof. If $L < 1$, then, as in the proof for the ratio test, we can choose $q < 1$ and find a corresponding N such that

$$b_n^* \equiv \sqrt[n]{|w_n|} < q < 1 \quad \text{for every } n > N.$$

From this we obtain

$$|w_n| < q^n < 1 \quad (n > N),$$

and the series $|w_N| + |w_{N+1}| + \cdots$ converges by comparison with the geometric series. Hence, the series $w_1 + w_2 + \cdots$ converges absolutely.

If $L > 1$, then $\sqrt[n]{|w_n|} > 1$ for all sufficiently large n . Hence $|w_n| > 1$ for those n , and divergence follows from Theorem 1 in the last section.

If $L = 1$, the test fails. This is illustrated by the harmonic series for which

$$\sqrt[n]{\frac{1}{n}} = \frac{1}{n^{1/n}} = \frac{1}{e^{(1/n)\ln n}} \rightarrow \frac{1}{e} = 1 \quad (n \rightarrow \infty)$$

because $(1/n) \ln n \rightarrow 0$, and the series (5), for which

$$\sqrt[n]{\frac{1}{n^2}} = \frac{1}{n^{2/n}} = \frac{1}{e^{(2/n)\ln n}} \rightarrow \frac{1}{e^0} = 1 \quad (n \rightarrow \infty)$$

because $(2/n) \ln n \rightarrow 0$ as $n \rightarrow \infty$.

Example 1. Test the series

$$\sum_{n=1}^{\infty} \frac{n^2}{2^n} = \frac{1}{2} + 1 + \frac{9}{8} + 1 + \frac{25}{32} + \dots$$

We have

$$w_n = \frac{n^2}{2^n}, \quad w_{n+1} = \frac{(n+1)^2}{2^{n+1}}, \quad \frac{w_{n+1}}{w_n} = \frac{(n+1)^2}{2n^2} \rightarrow \frac{1}{2} \quad (n \rightarrow \infty),$$

and convergence follows from the ratio test. We may also apply the root test:

$$\sqrt[n]{\frac{n^2}{2^n}} = \frac{n^{2/n}}{2} = \frac{e^{(2/n)\ln n}}{2} \rightarrow \frac{e^0}{2} = \frac{1}{2} \quad (n \rightarrow \infty).$$

This confirms our result.

Example 2. Is the series

$$\sum_{n=0}^{\infty} \frac{(3-4i)^n}{n!} = 1 + (3-4i) + \frac{1}{2!}(3-4i)^2 + \dots$$

convergent or divergent? We have

$$|w_n| = \frac{|3-4i|^n}{n!} = \frac{5^n}{n!}, \quad |w_{n+1}| = \frac{5^{n+1}}{(n+1)!}, \quad \left| \frac{w_{n+1}}{w_n} \right| = \frac{5}{n+1} \rightarrow 0$$

as $n \rightarrow \infty$, and convergence follows from the ratio test.

We finally mention that both the ratio test and the root test can be generalized to the case when the sequences $|w_{n+1}/w_n|$ and $\sqrt[n]{|w_n|}$, respectively, are not convergent.

Theorem 5 (Ratio test). If $w_n \neq 0$ for $n = 1, 2, \dots$, and if for every n greater than some number N

$$(6) \quad \left| \frac{w_{n+1}}{w_n} \right| \leq q,$$

where q is some fixed number less than 1, then the series

$$w_1 + w_2 + w_3 + \dots$$

converges absolutely. If

$$(7) \quad \left| \frac{w_{n+1}}{w_n} \right| \geq 1 \quad \text{for every } n > N,$$

the series diverges.

Proof. In the first part of the proof of Theorem 3, convergence followed from the existence of a number $q < 1$ such that (6) holds for all $n > N$. Hence, for the present theorem, convergence follows by the same argument. The

last statement of the theorem follows from the fact that, by (7), $|w_{n+1}| \geq |w_n|$ and Theorem 1 in the last section.

Example 3. In the series

$$1 + \frac{1}{2} + \frac{1}{8} + \frac{1}{16} + \frac{1}{64} + \frac{1}{128} + \frac{1}{512} + \cdots$$

the odd terms and the even terms each form a geometric series with ratio $1/8$. Convergence follows from Theorem 5, because the ratios of successive terms are

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{2}, \frac{1}{4}, \dots$$

We see that the sequence of these ratios does not converge. Hence Theorem 3 cannot be applied, and our example illustrates that Theorem 5 is more general than Theorem 3.

Theorem 6 (Root test). *If for every n greater than some number N*

$$(8) \quad \sqrt[n]{|w_n|} \leq q,$$

where q is some fixed number less than 1, then the series

$$w_1 + w_2 + w_3 + \cdots$$

converges absolutely. If for infinitely many n

$$(9) \quad \sqrt[n]{|w_n|} \geq 1,$$

the series diverges.

Proof. If (8) holds, then

$$|w_n| \leq q^n < 1 \quad (n > N)$$

and the series $|w_N| + |w_{N+1}| + \cdots$ converges by comparison with the geometric series. Hence, the series $w_1 + w_2 + \cdots$ converges absolutely. If (9) holds, then $|w_n| \geq 1$ for infinitely many n , and divergence follows from Theorem 1 in the last section.

In the last two theorems it is essential for convergence that $|w_{n+1}/w_n|$ and $\sqrt[n]{|w_n|}$, respectively, should be ultimately less than or equal to a definite number $q < 1$. It does not at all suffice for convergence that we should have

$$\left| \frac{w_{n+1}}{w_n} \right| < 1 \quad \text{or} \quad \sqrt[n]{|w_n|} < 1$$

for sufficiently large n . For example, in the case of the harmonic series (Ex. 4 in Sec. 13.2) we have

$$\frac{w_{n+1}}{w_n} = \frac{1/(n+1)}{1/n} = \frac{n}{n+1} < 1 \quad \text{and} \quad \sqrt[n]{\frac{1}{n}} < 1,$$

but the series diverges.

Problems

Are the following series convergent or divergent?

1. $1 - \frac{1}{3} + \frac{1}{9} - \frac{1}{27} + \dots$

2. $1 + \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \dots$

3. $1 + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{4}} + \dots$

4. $1 - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{4}} + \dots$

5. $\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \dots$

6. $\frac{1}{\sqrt{1 \cdot 2}} + \frac{1}{\sqrt{2 \cdot 3}} + \frac{1}{\sqrt{3 \cdot 4}} + \dots$

7. $1 + 10i + \frac{(10i)^2}{2!} + \frac{(10i)^3}{3!} + \dots$

8. $\frac{1}{\ln 2} + \frac{1}{\ln 3} + \frac{1}{\ln 4} + \dots$

9. $1 + \frac{i}{2} + \frac{i^2}{2^2} + \frac{i^3}{2^3} + \dots$

10. $1 + i + i^2 + i^3 + \dots$

11. $\sum_{n=1}^{\infty} \frac{n^n}{n!}$

12. $\sum_{n=0}^{\infty} \frac{(1+i)^n}{n!}$

13. $\sum_{n=0}^{\infty} \frac{(10+5i)^n}{n!}$

14. $\sum_{n=0}^{\infty} \frac{(3+i)^{2n}}{(2n)!}$

15. $\sum_{n=0}^{\infty} n \left(\frac{i}{2}\right)^n$

16. $\sum_{n=0}^{\infty} \left(\frac{4+3i}{6}\right)^n$

17. $\sum_{n=0}^{\infty} \left(\frac{3-4i}{4}\right)^n$

18. $\sum_{n=0}^{\infty} \frac{(-1)^n (10i)^{2n}}{(2n)!}$

19. $\sum_{n=1}^{\infty} \frac{i^n}{2^n n}$

20. If $|w_{n+1}/w_n| \leq q < 1$, so that the series $w_1 + w_2 + \dots$ converges by the ratio test (Theorem 5), show that the remainder $R_n = w_{n+1} + w_{n+2} + \dots$ satisfies $|R_n| \leq |w_{n+1}|/(1-q)$. Hint: use the fact that the ratio test is a comparison of the series $w_1 + w_2 + \dots$ with the geometric series.

21. Determine how many terms are needed to compute the sum s of the geometric series $1 + q + q^2 + \dots$ with an error less than 0.01, when $q = 0.25$, $q = 0.5$, $q = 0.9$.

Show that the following series converge. How many terms are needed for computing s with an error less than 0.01?

22. $s = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \dots$

23. $s = 1 - \frac{1}{3!} + \frac{1}{5!} - \frac{1}{7!} + \dots$

24. $s = 1 - \frac{1}{3} + \frac{1}{9} - \frac{1}{27} + \frac{1}{81} - \dots$

25. $s = 1 - \frac{1}{3} + \frac{1}{9} - \frac{1}{27} + \dots$

13.4 Operations on Series

We shall now consider some simple operations which are frequently used in working with series.

We start with the addition of series and show that two convergent series may be added term by term.

Theorem 1. *If the series $w_1 + w_2 + \dots$ and $z_1 + z_2 + \dots$ are convergent with the respective sums s and s^* , then the series*

$$(1) \quad \sum_{n=1}^{\infty} (w_n + z_n), \quad \sum_{n=1}^{\infty} (w_n - z_n), \quad \text{and} \quad \sum_{n=1}^{\infty} kw_n$$

where k is any constant, are convergent and have the sums $s + s^*$, $s - s^*$, and ks , respectively.

Proof. The partial sums of the two given series are

$$s_n = w_1 + \cdots + w_n, \quad s_n^* = z_1 + \cdots + z_n,$$

and by definition of convergence of a series,

$$\lim_{n \rightarrow \infty} s_n = s, \quad \lim_{n \rightarrow \infty} s_n^* = s^*.$$

Now the n th partial sum of the first series in (1) is

$$S_n = s_n + s_n^* = (w_1 + z_1) + \cdots + (w_n + z_n).$$

From this we obtain

$$\lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} (s_n + s_n^*) = \lim_{n \rightarrow \infty} s_n + \lim_{n \rightarrow \infty} s_n^* = s + s^*.$$

Hence this series converges and has the sum $s + s^*$. The other statements can be proved by a similar argument.

The next operation to be considered is that of inserting parentheses in a given series. This operation is called **grouping**.

For example, by grouping the terms of the series $w_1 + w_2 + w_3 + \cdots$ we may obtain the series $(w_1 + w_2) + (w_3 + w_4) + \cdots$ whose terms are $W_n = w_{2n-1} + w_{2n}$ where $n = 1, 2, \dots$

Of course, in the case of a *finite* series (that is, a series with only finitely many terms) we may insert parentheses without changing the sum of the series. However, this is not true in general for an infinite series. For instance, by grouping the divergent series

$$1 - 1 + 1 - 1 + \cdots \quad (\text{cf. Ex. 3 in Sec. 13.2})$$

we may obtain the convergent series

$$(1 - 1) + (1 - 1) + \cdots = 0 + 0 + \cdots$$

whose sum is zero.

We shall now prove that, in the case of a convergent series, grouping is permissible.

Theorem 2. *If a series converges, then insertion of parentheses yields a new convergent series having the same sum as the original series.*

Proof. Obviously, the partial sums of the new series are obtained by skipping certain partial sums of the given series. For example, the partial sums of the series

$$(w_1 + w_2 + w_3) + (w_4 + w_5 + w_6) + \cdots$$

are the partial sums s_3, s_6, s_9, \dots of the series $w_1 + w_2 + \cdots$. Now if the sequence of the partial sums s_n of the given series converges to s , it follows

immediately from the definition of convergence of a sequence that the new sequence obtained by skipping must also converge to s . This proves the theorem.

Example 1. By the Leibniz test (cf. Sec. 13.2) the series

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

converges. Let s be its sum. Then by Theorem 2,

$$(2) \quad \begin{aligned} \frac{1}{1 \cdot 2} + \frac{1}{3 \cdot 4} + \frac{1}{5 \cdot 6} + \dots &= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \dots \\ &= \sum_{m=1}^{\infty} \left(\frac{1}{2m-1} - \frac{1}{2m}\right) = s \end{aligned}$$

and, similarly,

$$(3) \quad \begin{aligned} \frac{1 \cdot 2 + 3 \cdot 4}{1 \cdot 2 \cdot 3 \cdot 4} + \frac{5 \cdot 6 + 7 \cdot 8}{5 \cdot 6 \cdot 7 \cdot 8} + \dots &= \left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8}\right) + \dots \\ &= \sum_{m=1}^{\infty} \left(\frac{1}{4m-3} - \frac{1}{4m-2} + \frac{1}{4m-1} - \frac{1}{4m}\right) = s. \end{aligned}$$

The last operation to be considered in this section is that of changing the order of the terms in a series.

Clearly, if a series is finite, the order of its terms may be changed without altering the sum. Also, we may change the order of finitely many terms of a given infinite series: if the given series is divergent, so will be the new series, and if the given series converges, the new series will be convergent with the same sum. This follows immediately from the definitions of convergence and divergence.

We may now ask what happens if we “change the order of infinitely many terms” of a series. Of course we must first define what we mean by such an operation. This can be done as follows.

A series

$$\sum_{n=1}^{\infty} w_n^* = w_1^* + w_2^* + \dots$$

is said to be a rearrangement of a series

$$\sum_{m=1}^{\infty} w_m = w_1 + w_2 + \dots$$

if there is a one-to-one correspondence between the indices n and m such that $w_n^* = w_m$ for corresponding indices.

For example, the series

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{3} + \frac{1}{6} + \frac{1}{5} + \dots$$

is a rearrangement of the harmonic series

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \dots$$

The following example illustrates that by rearranging a convergent series we may obtain a convergent series whose sum is different from that of the original series.

Example 2. We rearrange the series

$$s = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} - \dots$$

by writing first two positive terms, then one negative, then again two positive, etc., in the order of their occurrence. This yields the rearrangement

$$s^* = 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \frac{1}{13} + \dots$$

We want to show that the sums s and s^* are different. By inserting parentheses in the rearrangement we have

$$s^* = \left(1 + \frac{1}{3} - \frac{1}{2}\right) + \left(\frac{1}{5} + \frac{1}{7} - \frac{1}{4}\right) + \dots = \sum_{m=1}^{\infty} \left(\frac{1}{4m-3} + \frac{1}{4m-1} - \frac{1}{2m}\right).$$

On the other hand, by multiplying the series (2) by 1/2 and adding the resulting series and the series (3) term by term (cf. Theorem 1), we get

$$\begin{aligned} \frac{3s}{2} &= \sum_{m=1}^{\infty} \left(\frac{1}{4m-3} - \frac{1}{4m-2} + \frac{1}{4m-1} - \frac{1}{4m} + \frac{1/2}{2m-1} - \frac{1/2}{2m}\right) \\ &= \sum_{m=1}^{\infty} \left(\frac{1}{4m-3} + \frac{1}{4m-1} - \frac{1}{2m}\right) = s^*. \end{aligned}$$

Hence $s^* = 3s/2$.

We note that the series in Ex. 2 is not absolutely convergent. Let us prove that an absolutely convergent series can always be rearranged without altering the sum.

Theorem 3. If a series converges absolutely, then any of its rearrangements converges absolutely and has the same sum as the original series.

Proof. Let $w_1^* + w_2^* + w_3^* + \dots$ be any rearrangement of an absolutely convergent series $w_1 + w_2 + w_3 + \dots$. Then, since every w_m^* equals a w_n for appropriate n and no two m 's correspond to the same n we clearly have

$$\sum_{m=1}^n |w_m^*| \leq \sum_{k=1}^{\infty} |w_k| \quad \text{for every } n.$$

The sum on the left is the n th partial sum of the series $|w_1^*| + |w_2^*| + \dots$. Since these partial sums are nonnegative, the inequality shows that they form a bounded sequence. Since $|w_m^*| \geq 0$, the sequence is monotone increasing and, therefore, convergent (cf. Theorem 5 in Sec. 13.1). Hence the rearrangement $w_1^* + w_2^* + \dots$ converges absolutely. Let s^* be its sum, and let s be the sum of the original series. We show that $s^* = s$.

From the definition of convergence and from Theorem 4 in Sec. 13.2 when applied to the series $|w_1| + |w_2| + \dots$ it follows that, when $\epsilon > 0$ is given, we can find an N such that

$$(4) \quad (a) \quad |s_n - s| < \frac{\epsilon}{2} \quad (b) \quad |w_{n+1}| + \dots + |w_{n+p}| < \frac{\epsilon}{2}$$

for every $n > N$ and $p = 1, 2, \dots$; here s_n is the n th partial sum of the original series. Now for sufficiently large m , the m th partial sum s_m^* of the rearrangement will contain all the terms w_1, \dots, w_n ($n > N$ and fixed) and perhaps some more terms w_r ($r > n$) of the original series. Hence s_m^* will be of the form

$$(5) \quad s_m^* = s_n + A_{mn}$$

where A_{mn} is the sum of those additional terms. Let $n + p$ be the greatest subscript of the terms in A_{mn} . Then by (4b), since $n > N$,

$$|A_{mn}| \leq |w_{n+1}| + \dots + |w_{n+p}| < \frac{\epsilon}{2}.$$

From this and (5) it follows that

$$|s_m^* - s_n| = |A_{mn}| < \frac{\epsilon}{2}.$$

Using (4a) and the triangle inequality, we thus obtain

$$|s_m^* - s| = |(s_m^* - s_n) + (s_n - s)| \leq |s_m^* - s_n| + |s_n - s| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

for every sufficiently large m . Hence, s_m^* converges to s and, therefore, $s^* = s$. This completes the proof.

13.5 Power Series

We have made the definitions for series of *constant* terms. If the terms of a series are *variable*, say, functions of a variable z , they assume definite values when z is given a fixed value and then all those definitions apply. It is obvious that in the case of such a series of functions of z the partial sums, the remainders, and the sum will be functions of z . Usually such a series will be convergent for some values of z , for example, for all z in some region, and divergent for the other values of z .

In complex analysis, the most important series with variable terms are power series. A power series⁵ in powers of $z - a$ is an infinite series of the form

$$(1) \quad \sum_{m=0}^{\infty} c_m(z - a)^m = c_0 + c_1(z - a) + c_2(z - a)^2 + \dots$$

⁵ It should be noted that the term "power series" alone usually refers to series of the form (1), including the particular case (2), but does not include series of negative powers of z such as $c_1 z^{-1} + c_2 z^{-2} + \dots$ or series involving fractional powers of z .

where z is a variable, c_0, c_1, \dots are constants, called the coefficients, and a is a constant, called the center of the series.

If $a = 0$, we obtain as a particular case a power series in powers of z :

$$(2) \quad \sum_{m=0}^{\infty} c_m z^m = c_0 + c_1 z + c_2 z^2 + \dots$$

The convergence behavior of a power series can be characterized in a very simple way. Let us start with three examples, which will turn out to be typical.

Example 1. The geometric series

$$\sum_{m=0}^{\infty} z^m = 1 + z + z^2 + \dots$$

converges absolutely when $|z| < 1$ and diverges when $|z| \geq 1$ (cf. Theorem 2 in Sec. 13.3).

Example 2. The power series

$$\sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots$$

is absolutely convergent for each (finite) z , as follows from the ratio test, because

$$\left| \frac{z^{n+1}/(n+1)!}{z^n/n!} \right| = \frac{|z|}{n+1} \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty \quad (z \text{ fixed}).$$

Example 3. The power series

$$\sum_{n=0}^{\infty} n! z^n = 1 + z + 2z^2 + 6z^3 + \dots$$

converges only at the center $z = 0$, but diverges for any $z \neq 0$. In fact, this follows from the ratio test because

$$\left| \frac{(n+1)! z^{n+1}}{n! z^n} \right| = (n+1)|z| \rightarrow \infty \quad \text{as} \quad n \rightarrow \infty \quad (z \text{ fixed and } \neq 0).$$

The power series (1) converges when $z = a$, because then $z - a = 0$ and the series reduces to the single term c_0 . Example 3 illustrates that in some cases this may be the only value of z for which the series converges. However, if the series converges for some $z_0 \neq a$, then it converges for every z whose distance from the center is less than that of z_0 . In fact, the following theorem holds.

Theorem 1. If the power series (1) converges for $z = z_0$, it converges absolutely for every z for which $|z - a| < |z_0 - a|$, that is, for each z within the circle through z_0 about a .

Proof. Since the series (1) converges for z_0 , Theorem 1 in Sec. 13.2 shows that

$$c_n(z_0 - a)^n \rightarrow 0 \quad \text{as} \quad n \rightarrow \infty.$$

This implies that for $z = z_0$ the terms of (1) are bounded, say,

$$\bullet \quad |c_n(z_0 - a)^n| < M \quad \text{for every } n = 0, 1, \dots$$

From this we obtain

$$|c_n(z - a)^n| = \left| c_n(z_0 - a)^n \left(\frac{z - a}{z_0 - a} \right)^n \right| < M \left| \frac{z - a}{z_0 - a} \right|^n$$

and, therefore,

$$(3) \quad \sum_{n=0}^{\infty} |c_n(z - a)^n| < \sum_{n=0}^{\infty} M \left| \frac{z - a}{z_0 - a} \right|^n = M \sum_{n=0}^{\infty} \left| \frac{z - a}{z_0 - a} \right|^n$$

From the assumption $|z - a| < |z_0 - a|$ we have

$$\left| \frac{z - a}{z_0 - a} \right| < 1,$$

and the series on the right-hand side of (3) is a geometric series which converges. Hence, the series on the left-hand side of (3) converges, and the series (1) converges absolutely when $|z - a| < |z_0 - a|$. This completes the proof.

Examples 2 and 3 illustrate that a power series may converge for all z or only for $z = a$. Let us exclude these two cases from our consideration for a moment. Then, if a power series (1) is given, we may consider all the points z in the complex plane for which the series converges. Let R be the smallest real number so that the distance of any of these points from the center a is at most equal to R . (For instance, in Ex. 1, $R = 1$.) Then from Theorem 1 it follows that the series converges for all z within the circle of radius R about the center a , that is, for all z for which

$$(4) \quad |z - a| < R$$

and diverges for all z for which

$$|z - a| > R.$$

The circle

$$|z - a| = R$$

is called the **circle of convergence**, and its radius R is called the **radius of convergence** of (1). Cf. Fig. 339.

At the points on the circle of convergence the series may converge or diverge. For instance, in Ex. 1 we have $R = 1$, and the series diverges at each point of the circle of convergence $|z| = 1$. The series

$$\sum_{n=1}^{\infty} \frac{z^n}{n} = z + \frac{z^2}{2} + \frac{z^3}{3} + \dots$$

converges for $|z| < 1$ and diverges for $|z| > 1$, as follows by the ratio test. Thus $R = 1$. At $z = 1$ it becomes the harmonic series and diverges. At $z = -1$ it becomes the series in Ex. 5, Sec. 13.2, and converges. This illustrates that a series may converge at some points on the circle of convergence, while at others it may diverge.

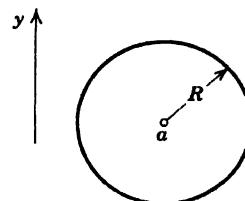


Fig. 339. Circle of convergence.

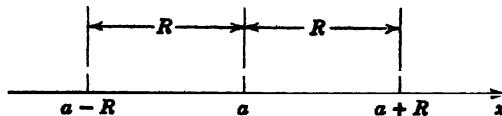


Fig. 340. Interval of convergence of a real power series.

Clearly, if we consider a *real* power series (1), that is if the variable $z = x$, the center, and the coefficients are real, then (4) represents an interval of length $2R$ with midpoint at a on the x -axis, the so-called **interval of convergence** (Fig. 340).

If the series (1) converges for all z (as in Ex. 2), then we set

$$R = \infty \quad (\text{and } 1/R = 0);$$

if it converges only at the center $z = a$ (as in Ex. 3), then we set

$$R = 0 \quad (\text{and } 1/R = \infty).$$

Using these conventions, the radius of convergence R of the power series (1) may be determined from the coefficients of the series as follows.

Theorem 2. *If the sequence $\sqrt[n]{|c_n|}$, $n = 1, 2, \dots$, converges with the limit L , then the radius of convergence R of the power series (1) is*

$$(5a) \quad R = \frac{1}{L} \quad (\text{Cauchy-Hadamard formula}),^6$$

including the case $L = 0$ where $R = \infty$ and the series (1) converges for all z .

If that sequence does not converge but is bounded, then

$$(5b) \quad R = \frac{1}{l}$$

where l is the greatest of the limit points of the sequence.

If that sequence is not bounded, then $R = 0$, and the series converges only for $z = a$.

Proof. If

$$\lim_{n \rightarrow \infty} \sqrt[n]{|c_n|} = L \neq 0,$$

then

$$\lim_{n \rightarrow \infty} \sqrt[n]{|c_n(z - a)^n|} = |z - a| \lim_{n \rightarrow \infty} \sqrt[n]{|c_n|} = |z - a|L.$$

Since the terms of the series (1) are $w_n = c_n(z - a)^n$, the root test (Sec. 13.3) shows that the series converges absolutely when

$$|z - a|L < 1 \quad \text{or} \quad |z - a| < \frac{1}{L} = R$$

⁶ Named after the French mathematicians, A. L. CAUCHY (cf. the footnote in Sec. 1.13) and JACQUES HADAMARD (1865–1964).

but diverges when

$$|z - a|L > 1 \quad \text{or} \quad |z - a| > \frac{1}{L} = R.$$

If

$$\lim_{n \rightarrow \infty} \sqrt[n]{|c_n|} = L = 0,$$

it follows from the definition of limit that for any given $\epsilon > 0$, for example, for $\epsilon = 1/(2|z_1 - a|)$ with arbitrary fixed z_1 , we can find an N such that

$$\sqrt[n]{|c_n|} < \frac{1}{2|z_1 - a|} \quad \text{for every } n > N.$$

From this we immediately obtain

$$|c_n| < \frac{1}{(2|z_1 - a|)^n} \quad \text{or} \quad |c_n(z_1 - a)^n| < \frac{1}{2^n}.$$

Now, since $\sum 2^{-n}$ converges, the comparison test (cf. Sec. 13.3) shows that the series (1) converges absolutely for $z = z_1$. Since z_1 is arbitrary, this means absolute convergence for every finite z , and the proof of the statement involving (5a) is complete.

We next prove the statement involving (5b). The existence of l follows from the Bolzano-Weierstrass theorem (Sec. 13.1), and since $\sqrt[n]{|c_n|} \geq 0$, we clearly have $l \geq 0$. From the definition of a limit point it follows that, an $\epsilon > 0$ being given,

$$l - \epsilon < \sqrt[n]{|c_n|} < l + \epsilon \quad \text{for infinitely many } n.$$

By multiplying this by the positive quantity $|z - a|$ we have

$$(6) \quad |z - a|(l - \epsilon) < \sqrt[n]{|c_n(z - a)^n|}$$

and

$$(7) \quad \sqrt[n]{|c_n(z - a)^n|} < |z - a|(l + \epsilon).$$

The inequality (7) holds even for all sufficiently large n , say, for $n > N$, because l is the greatest limit point and, therefore, at most finitely many terms can be greater than the expression on the right. We prove that for

$$(8) \quad |z - a| < \frac{1}{l}$$

convergence follows from (7). In fact, if we choose

$$\epsilon = \frac{1 - l|z - a|}{2|z - a|},$$

then because of (8) we have $\epsilon > 0$, and (7) becomes

$$\sqrt[n]{|c_n(z - a)^n|} < \frac{1 + l|z - a|}{2}. \quad (n > N)$$

From (8) we see that the expression on the right is less than 1, and convergence follows by means of the root test (Theorem 6, Sec. 13.3). On the other hand, when

$$|z - a| > \frac{1}{l},$$

then, choosing

$$\epsilon = \frac{l|z - a| - 1}{2|z - a|},$$

we have $\epsilon > 0$, and (6) becomes

$$\sqrt[n]{|c_n(z - a)^n|} > \frac{|z - a| l + 1}{2} > 1.$$

Hence, by Theorem 6 in Sec. 13.3 the series diverges for those z . This proves the statement involving (5b).

Finally if the sequence $\sqrt[n]{|c_n|}$ is not bounded, then, by definition, any K being given,

$$\sqrt[n]{|c_n|} > K \quad \text{for infinitely many } n.$$

Choosing $K = 1/|z - a|$ where $z \neq 0$, we see that this becomes

$$\sqrt[n]{|c_n|} > \frac{1}{|z - a|} \quad \text{or} \quad \sqrt[n]{|c_n(z - a)^n|} > 1,$$

and divergence follows from Theorem 6 in Sec. 13.3. This completes the proof of Theorem 2.

We shall now consider the operations of addition and multiplication of power series.

From Theorem 1, Sec. 13.4, it is clear that two power series may be added term by term for any z for which both series are convergent.

Let us consider the term-by-term multiplication of two power series

$$(9) \quad \sum_{k=0}^{\infty} a_k z^k = a_0 + a_1 z + \cdots \quad \text{and} \quad \sum_{m=0}^{\infty} c_m z^m = c_0 + c_1 z + \cdots.$$

If we multiply each term of the first series by each term of the second series and collect products of like powers of z , we obtain

$$(10) \quad \begin{aligned} a_0 c_0 + (a_0 c_1 + a_1 c_0) z + (a_0 c_2 + a_1 c_1 + a_2 c_0) z^2 + \cdots \\ = \sum_{n=0}^{\infty} (a_0 c_n + a_1 c_{n-1} + \cdots + a_n c_0) z^n. \end{aligned}$$

This series is called the **Cauchy product** of the series (9).

Theorem 3. *The Cauchy product of the two power series (9) is absolutely convergent for each z within the circle of convergence of each of the series (9).*

If the series have the sums $g(z)$ and $h(z)$, respectively, the Cauchy product has the sum

$$(11) \quad s(z) = g(z)h(z).$$

Proof. The general term of the product series (10) is

$$p_n = (a_0c_n + a_1c_{n-1} + \cdots + a_nc_0)z^n.$$

Now, by the generalized triangle inequality (3) in Sec. 10.1,

$$|p_0| + |p_1| = |a_0c_0| + |(a_0c_1 + a_1c_0)z| \leq (|a_0| + |a_1z|)(|c_0| + |c_1z|),$$

$$|p_0| + |p_1| + |p_2| \leq (|a_0| + |a_1z| + |a_2z^2|)(|c_0| + |c_1z| + |c_2z^2|),$$

as can be readily verified by performing the multiplication on the right; in general,

$$\begin{aligned} |p_0| + |p_1| + \cdots + |p_n| \\ \leq (|a_0| + |a_1z| + \cdots + |a_nz^n|)(|c_0| + |c_1z| + \cdots + |c_nz^n|). \end{aligned}$$

If z lies within the circle of convergence of each of the series (9), the sequence of the expressions on the right is bounded, and so is the sequence of the partial sums of the series $|p_0| + |p_1| + \cdots$. Since $|p_n| \geq 0$, that sequence is monotone increasing, and from Theorem 5 in Sec. 13.1 it follows that it converges. Hence that series is convergent, and the product series (10) converges absolutely.

We prove (11). By Theorem 3 in Sec. 13.4 any rearrangement of (10) is absolutely convergent for precisely those z and has the same sum as (10). We consider the particular rearrangement $p_0^* + p_1^* + \cdots$ where (Fig. 341)

$$\begin{aligned} p_n^* = (a_nc_0 + a_0c_n)z^n + (a_nc_1 + a_1c_n)z^{n+1} + \cdots \\ + (a_nc_{n-1} + a_{n-1}c_n)z^{2n-1} + a_nc_nz^{2n}. \end{aligned}$$

Obviously,

$$a_0c_0 = p_0^*, \quad (a_0 + a_1z)(c_0 + c_1z) = p_0^* + p_1^*$$

and, in general,

$$(a_0 + a_1z + \cdots + a_nz^n)(c_0 + c_1z + \cdots + c_nz^n) = p_0^* + p_1^* + \cdots + p_n^*.$$

By letting n approach infinity we obtain (11), and Theorem 3 is proved.

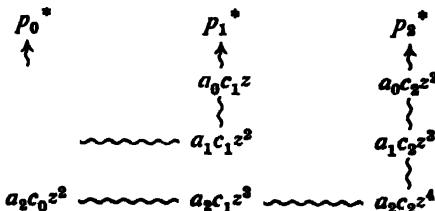


Fig. 341. Proof of Theorem 3.

Example 4. The geometric series $1 + z + z^2 + \dots$ has the sum $1/(1 - z)$ when $|z| < 1$ (cf. Theorem 2 in Sec. 13.3), and from Theorem 3 it follows that

$$\begin{aligned} \left(\frac{1}{1-z}\right)^2 &= \sum_{k=0}^{\infty} z^k \sum_{m=0}^{\infty} z^m = (1 + z + z^2 + \dots)(1 + z + z^2 + \dots) \\ &= 1 + 2z + 3z^2 + \dots = \sum_{n=0}^{\infty} (n+1)z^n \quad (|z| < 1). \end{aligned}$$

Problems

- Show that if the sequence $|c_{n+1}/c_n|$, $n = 1, 2, \dots$, is convergent with the limit L , then the radius of convergence R of the power series (1) is $R = 1/L$ when $L > 0$ and is $R = \infty$ when $L = 0$.
- Show that if the radius of convergence of (2) is R , then the radius of convergence of $\sum c_n z^{2n}$ is \sqrt{R} .

Find the radius of convergence of the following series.

- | | | |
|--|---|---|
| 3. $\sum_{n=0}^{\infty} (z - 2i)^n$ | 4. $\sum_{n=0}^{\infty} \frac{(z - 2)^n}{2^n}$ | 5. $\sum_{n=0}^{\infty} n \left(\frac{z}{3}\right)^n$ |
| 6. $\sum_{n=0}^{\infty} \frac{(2z)^n}{n!}$ | 7. $\sum_{n=0}^{\infty} \frac{z^{2n}}{n!}$ | 8. $\sum_{n=1}^{\infty} \frac{(-1)^n}{n} z^n$ |
| 9. $\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} z^n$ | 10. $\sum_{n=0}^{\infty} \frac{n^2}{2^n} z^n$ | 11. $\sum_{n=0}^{\infty} \frac{(2n)!}{(n!)^2} z^n$ |
| 12. $\sum_{n=0}^{\infty} \frac{z^n}{n^n}$ | 13. $\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n}$ | 14. $\sum_{n=1}^{\infty} \frac{z^n}{n^2}$ |
| 15. $\sum_{n=0}^{\infty} 6^n (z - i)^n$ | 16. $\sum_{n=0}^{\infty} (n!)^2 z^n$ | 17. $\sum_{n=0}^{\infty} 3^{2n} z^n$ |
| 18. $\sum_{n=0}^{\infty} \left(\frac{\pi i}{2}\right)^n z^n$ | 19. $\sum_{n=0}^{\infty} \frac{(n!)^2}{(2n)!} z^n$ | 20. $\sum_{n=0}^{\infty} n^{2n} z^n$ |

13.6 Functions Represented by Power Series

Power series are basic in complex analysis, and the main objective of this section is to show the reasons for this fact.

Let $\sum_{n=0}^{\infty} c_n z^n$ be an arbitrary power series with nonzero radius of convergence R . Then the sum of this series is a function of z , say, $f(z)$, and we write

$$(1) \quad f(z) = \sum_{n=0}^{\infty} c_n z^n = c_0 + c_1 z + c_2 z^2 + \dots \quad (|z| < R).$$

We say that $f(z)$ is represented by the power series or that it is developed in the power series. For example, the geometric series in Ex. 1, Sec. 13.5,

represents the function $f(z) = 1/(1 - z)$ in the interior of the unit circle $|z| = 1$. This follows from Theorem 2 in Sec. 13.3.

We shall now show that each power series with a nonzero radius of convergence represents an analytic function. In Sec. 14.1 we shall prove that, conversely, every analytic function can be represented by power series. For these reasons power series play a basic role in connection with analytic functions and their practical applications.

We subdivide our consideration into several steps, starting with the following theorem.

Theorem 1. *The function $f(z)$ in (1) is continuous at $z = 0$.*

Proof. We have to show that

$$(2) \quad \lim_{z \rightarrow 0} f(z) = f(0) = c_0.$$

We choose an arbitrary positive number $r < R$. Since the series in (1) is absolutely convergent in the disk $|z| < R$, it follows that the series

$$\sum_{n=1}^{\infty} |c_n| r^{n-1} = \frac{1}{r} \sum_{n=1}^{\infty} |c_n| r^n \quad (0 < r < R)$$

is convergent. Let K denote its sum. Then

$$|f(z) - c_0| = \left| z \sum_{n=1}^{\infty} c_n z^{n-1} \right| \leq |z| \sum_{n=1}^{\infty} |c_n| |z|^{n-1} \leq |z| K \quad (0 < |z| \leq r).$$

An $\epsilon > 0$ being given, $|f(z) - c_0| < \epsilon$ for all $|z| < \delta$ where δ is a positive real number less than both r and ϵ/K . By the definition of a limit, this means that (2) holds, and the proof is complete.

Next we consider the question of **uniqueness** and show that *the same function $f(z)$ cannot be represented by two different power series with the same center*. If $f(z)$ can be developed in a power series with center a , the development is unique. This important fact is frequently used in real and complex analysis. We may formulate it as follows (assuming that $a = 0$, without loss of generality).

Theorem 2 (Identity theorem for power series). *Suppose that*

$$\sum_{n=0}^{\infty} a_n z^n \quad \text{and} \quad \sum_{n=0}^{\infty} b_n z^n$$

are power series which are convergent for $|z| < R$ where R is positive, and have the same sum for all these z . Then the series are identical, that is,

$$(3) \quad a_n = b_n \quad \text{for all } n = 0, 1, \dots.$$

Proof. We proceed by induction. By assumption,

$$(4) \quad a_0 + a_1 z + a_2 z^2 + \dots = b_0 + b_1 z + b_2 z^2 + \dots \quad (|z| < R).$$

Let z approach zero. Then, by Theorem 1, $a_0 = b_0$. Assume that $a_n = b_n$ for $n = 0, 1, \dots, m$. Then, by omitting the first $m + 1$ terms on both sides of (4) and dividing by z^{m+1} ($\neq 0$), we obtain

$$a_{m+1} + a_{m+2}z + a_{m+3}z^2 + \dots : b_{m+1} + b_{m+2}z + b_{m+3}z^2 + \dots$$

By Theorem 1, each of these power series represents a function which is continuous at $z = 0$. Hence $a_{m+1} = b_{m+1}$. This completes the proof.

We shall now consider developments around different centers. Let us show that the function $f(z)$ represented by the power series (1) (with nonzero radius of convergence R) can also be represented by a power series with center $z = b$, where b is any point whose distance from the origin is less than R .

To derive such a power series, we set $z - b = Z$. Then $z = b + Z$, and (1) takes the form

$$(5) \quad f(z) = \sum_{n=0}^{\infty} c_n z^n = \sum_{n=0}^{\infty} c_n (b + Z)^n \\ = c_0 + c_1(b + Z) + c_2(b^2 + 2bZ + Z^2) + \dots$$

We consider the corresponding series

$$(6) \quad \sum_{n=0}^{\infty} |c_n| (|b| + |Z|)^n = |c_0| + |c_1| (|b| + |Z|) + |c_2| (|b| + |Z|)^2 + \dots$$

Since the original series converges absolutely when $|z| < R$, it follows that if

$$|b| + |Z| < R,$$

that is,

$$|Z| = |z - b| < R - |b|,$$

the series (6) is convergent. The corresponding values of z lie in the open disk D^* in Fig. 342. For these values of z we may thus arrange the series on

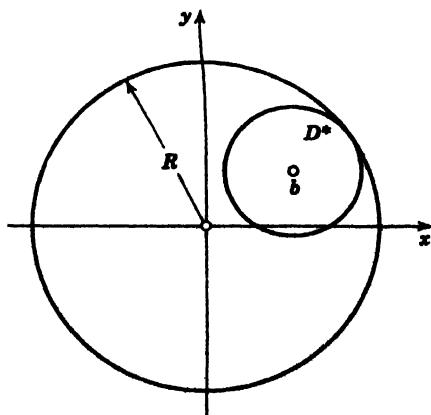


Fig. 342. Development around different centers.

the right-hand side of (5) in powers of $Z = z - b$ without altering its sum. In this way we obtain a representation of the form

$$(7) \quad f(z) = \sum_{n=0}^{\infty} a_n(z - b)^n,$$

valid (at least) in the disk

$$|z - b| < R - |b|.$$

We shall see later (in Sec. 14.1) that in general the radius of convergence of (7) will be greater than $R - |b|$, so that (7) will yield a "continuation" of the function $f(z)$ in (1) to points outside the disk $|z| < R$. This process of continuing a function, at first given by a power series in a region of convergence $|z| < R$, beyond that region is called **analytic continuation**.

By direct calculation it follows that the coefficients a_n can be represented in terms of the coefficients c_n of the original representation (1) in the form

$$(8) \quad a_n = \sum_{m=0}^{\infty} \binom{m+n}{n} c_{m+n} b^m \quad \text{where} \quad \binom{m+n}{n} = \frac{(m+n)!}{n! m!}$$

The practical application of (8) will be difficult in general. We shall see later (in Sec. 14.3) that in the case where $f(z)$ in (1) is a known function, there are various other methods for determining the coefficients of the corresponding series (7).

Let us now consider termwise differentiation and integration of power series. By differentiating the series $c_0 + c_1 z + c_2 z^2 + \dots$ we obtain the series

$$(9) \quad \sum_{n=1}^{\infty} n c_n z^{n-1} = c_1 + 2c_2 z + 3c_3 z^2 + \dots$$

which is called the *derived series* of the given series.

Theorem 3. *The derived series of a power series has the same radius of convergence as the original series.*

Proof. Let $nc_n = c_n^*$. Then $\sqrt[n]{|c_n^*|} = \sqrt[n]{n\sqrt[n]{|c_n|}}$. Since $\sqrt[n]{n} \rightarrow 1$ as $n \rightarrow \infty$, it follows that the sequences $\sqrt[n]{|c_n^*|}$ and $\sqrt[n]{|c_n|}$ either both converge with the same limit or both diverge. If they diverge, they are both unbounded or bounded, and in the latter case their greatest limit points are the same. From this and Theorem 2 in the last section the statement of our present theorem follows.

Theorem 4. *The series*

$$\sum_{n=0}^{\infty} \frac{c_n}{n+1} z^{n+1} = c_0 z + \frac{c_1}{2} z^2 + \frac{c_2}{3} z^3 + \dots$$

obtained by integrating the series $c_0 + c_1 z + c_2 z^2 + \dots$ term by term has the same radius of convergence as the original series.

The proof is similar to that of Theorem 3.

Example 1. The power series

$$\sum_{n=1}^{\infty} (n+1)z^n$$

has the radius of convergence $R = 1$, since by integrating term by term we obtain the geometric series $\sum_{n=1}^{\infty} z^{n+1}$.

Power series represent analytic functions. More precisely:

Theorem 5. *A power series with a nonzero radius of convergence R represents an analytic function at every point interior to its circle of convergence. The derivatives of this function are obtained by differentiating the original series term by term; all the series thus obtained have the same radius of convergence as the original series.*

Proof. We consider the representation

$$f(z) = \sum_{n=0}^{\infty} c_n z^n,$$

assuming that the radius of convergence R is not zero. Then we may represent $f(z)$ in the form (7), and from Theorem 1 it follows that $f(z)$ is continuous at the center b . Since b is any point in the disk $|z| < R$, the function $f(z)$ is continuous everywhere in this disk. From (7) we obtain $f(b) = a_0$ and, therefore,

$$\frac{f(z) - f(b)}{z - b} = a_1 + a_2(z - b) + a_3(z - b)^2 + \dots$$

According to Theorem 1, the function represented by the power series on the right is continuous at b . Hence by (8), with $n = 1$, we have

$$(10) \quad \begin{aligned} f'(b) &= \lim_{z \rightarrow b} \frac{f(z) - f(b)}{z - b} = a_1 = \sum_{m=0}^{\infty} (m+1)c_{m+1}b^m \\ &= \sum_{k=1}^{\infty} k c_k b^{k-1}. \end{aligned}$$

That is, the first derivative of $f(z)$ exists at any point $z = b$ in the disk $|z| < R$ and can be obtained by differentiating the original series term by term; indeed, the series on the right has the radius of convergence R , as was proved in Theorem 3. Hence $f(z)$ is analytic in the disk $|z| < R$, and the proof is complete.

Since the derivative

$$f'(z) = \sum_{n=1}^{\infty} n c_n z^{n-1}$$

of $f(z)$ is now represented by a power series, we may apply Theorem 5 to $f'(z)$ and conclude that $f''(z)$ exists in the disk $|z| < R$, and

$$f''(z) = \sum_{n=2}^{\infty} n(n-1)c_n z^{n-2}.$$

More generally, the m th derivative $f^{(m)}(z)$ of $f(z)$ exists in that disk, and

$$(11) \quad f^{(m)}(z) = \sum_{n=m}^{\infty} n(n-1)\cdots(n-m+1)c_n z^{n-m}.$$

In the next section we shall see that *every analytic function can be represented by power series*.

Problems

1. Using Theorem 2, show that if $f(z)$ in (1) is an odd function, then $c_n = 0$ for $n = 0, 2, 4, \dots$. Give examples.
2. Show that if $f(z)$ in (1) is even, then $c_n = 0$ for $n = 1, 3, 5, \dots$.
3. Show that $f(z) = \sum_{n=0}^{\infty} z^n/n!$ is neither even nor odd.
4. Verify Theorems 3 and 4 for the geometric series.
5. Verify Theorems 3 and 4 for the series in Prob. 3.

Using the geometric series and Theorems 3 and 4, find the radius of convergence of the following series.

$$6. \sum_{n=0}^{\infty} n(z-i)^n$$

$$7. \sum_{n=k}^{\infty} \binom{n}{k} z^n$$

$$8. \sum_{n=k}^{\infty} \binom{n}{k} \left(\frac{z}{3}\right)^n$$

$$9. \sum_{n=1}^{\infty} \frac{(-1)^n}{n} (z+1)^n$$

$$10. \sum_{n=2}^{\infty} \frac{z^n}{2^n n(n-1)}$$

$$11. \sum_{n=1}^{\infty} \frac{(z-2i)^n}{3^n n}$$

12. Explain the use of Theorems 2 and 5 in connection with the power series method in Sec. 3.1.

Using Theorem 5, prove the following propositions.

$$13. f(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \quad \text{satisfies} \quad f'(z) = f(z).$$

$$14. f(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} \quad \text{satisfies} \quad f''(z) + f(z) = 0.$$

$$15. f(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} \quad \text{satisfies} \quad f''(z) + f(z) = 0.$$

Chapter 14

Taylor and Laurent Series

In the last section we have seen that power series with nonzero radius of convergence represent analytic functions. We shall now see that *every* analytic function can be represented by power series. These representations are called Taylor series and are quite similar to the familiar Taylor series of real functions. In fact, by replacing the real variable in the latter series by a complex variable we may “*continue*” real functions analytically to the complex domain.

The last part of the chapter is devoted to the representation of analytic functions by Laurent series. These are series involving positive and negative integral powers of the independent variable. They are useful for evaluating complex and real integrals, as we shall see in the next chapter.

Prerequisites for this chapter: Chaps. 10, 12, 13.

Sections which may be omitted in a shorter course: 14.4, 14.6.

References: Appendix 1, Part G.

Answers to problems: Appendix 2.

✓14.1 *Taylor Series*

The familiar Taylor expansion is an effective tool in real calculus and its applications. We shall see that in complex analysis the Taylor expansion, which is a generalization of the aforementioned expansion, is even more important.

We consider a function $f(z)$ which is analytic in a neighborhood of a point $z = a$. Let C be a circle which lies in this neighborhood and has the center a . Then we may apply Cauchy's integral formula (1), Sec. 12.5; writing z and z^* in place of z_0 and z , we have

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(z^*)}{z^* - z} dz^*$$

where z is an arbitrary fixed point inside C and z^* is the complex variable of integration. Now,

$$(2) \quad \frac{1}{z^* - z} = \frac{1}{z^* - a - (z - a)} = \frac{1}{(z^* - a)\left(1 - \frac{z - a}{z^* - a}\right)}.$$

We note that since z^* is on C while z is inside C ,

$$(3) \quad \left| \frac{z - a}{z^* - a} \right| < 1.$$

From the geometric progression

$$1 + q + q^2 + \cdots + q^n = \frac{1 - q^{n+1}}{1 - q} \quad (q \neq 1)$$

we obtain the relation

$$\frac{1}{1 - q} = 1 + q + \cdots + q^n + \frac{q^{n+1}}{1 - q}.$$

By setting $q = (z - a)/(z^* - a)$ we find

$$\begin{aligned} 1 - [(z - a)/(z^* - a)] &= 1 + \frac{z - a}{z^* - a} + \left(\frac{z - a}{z^* - a}\right)^2 + \cdots + \left(\frac{z - a}{z^* - a}\right)^n \\ &\quad + \frac{[(z - a)/(z^* - a)]^{n+1}}{(z^* - z)/(z^* - a)}. \end{aligned}$$

We insert this into (2), and then (2) into (1). Since z and a are constant, we may take the powers of $z - a$ out from under the integral sign, and (1) takes the form

$$(4) \quad \begin{aligned} f(z) &= \frac{1}{2\pi i} \int_C \frac{f(z^*)}{z^* - a} dz^* + \frac{z - a}{2\pi i} \int_C \frac{f(z^*)}{(z^* - a)^2} dz^* + \cdots \\ &\quad + \frac{(z - a)^n}{2\pi i} \int_C \frac{f(z^*)}{(z^* - a)^{n+1}} dz^* + R_n(z) \end{aligned}$$

where the last term is given by the formula

$$(5) \quad R_n(z) = \frac{(z - a)^{n+1}}{2\pi i} \int_C \frac{f(z^*)}{(z^* - a)^{n+1}(z^* - z)} dz^*.$$

Using (1), Sec. 12.6, we may write this expansion in the form

$$(6) \quad \begin{aligned} f(z) &= f(a) + \frac{z - a}{1!} f'(a) + \frac{(z - a)^2}{2!} f''(a) + \cdots + \frac{(z - a)^n}{n!} f^{(n)}(a) \\ &\quad + R_n(z). \end{aligned}$$

This representation is called **Taylor's¹ formula**. $R_n(z)$ is called the *remainder*.

¹ BROOK TAYLOR (1685–1731), English mathematician, who introduced this formula for functions of a real variable.

Since the analytic function $f(z)$ has derivatives of all orders, we may take n in (6) as large as we please. If we let n approach infinity, we obtain from (6) the power series

$$(7) \quad f(z) = \sum_{m=0}^{\infty} \frac{f^{(m)}(a)}{m!} (z - a)^m.$$

This series is called the **Taylor series** of $f(z)$ with center at a . The particular case where $a = 0$ is called the **Maclaurin² series** of $f(z)$.

Clearly, the series (7) will converge and represent $f(z)$ if, and only if,

$$(8) \quad \lim_{n \rightarrow \infty} R_n(z) = 0.$$

To prove (8), we consider (5). Since z^* is on C while z is inside C , we have $|z^* - z| > 0$. Since $f(z)$ is analytic inside C and on C , it follows that the absolute value of $f'(z^*)/(z^* - z)$ is bounded, say,

$$\left| \frac{f(z^*)}{z^* - z} \right| < \bar{M}$$

for all z^* on C . Let r be the radius of C . Then C has the length $2\pi r$, and $|z^* - a| = r$ for all z^* on C . Hence, by applying (4) in Sec. 12.2 to (5) in the current section we obtain

$$\begin{aligned} |R_n| &= \frac{|z - a|^{n+1}}{2\pi} \left| \int_C \frac{f(z^*)}{(z^* - a)^{n+1}(z^* - z)} dz^* \right| \\ &< \frac{|z - a|^{n+1}}{2\pi} \bar{M} \frac{1}{r^{n+1}} 2\pi r = \bar{M}r \left| \frac{z - a}{r} \right|^{n+1}. \end{aligned}$$

If we let n approach infinity, it follows from (3) that the expression on the right approaches zero. This proves (8) for all z inside C . Since, by Theorem 2 in Sec. 13.6, the representation of $f(z)$ in the form (7) is unique in the sense that (7) is the only power series with center at a which represents the given function $f(z)$, we may sum up our result as follows.

Taylor's theorem. *Let $f(z)$ be analytic in a domain D and let $z = a$ be any point in D . Then there exists precisely one power series with center at a which represents $f(z)$; this series is of the form*

$$(9) \quad f(z) = \sum_{n=0}^{\infty} b_n(z - a)^n$$

where

$$b_n = \frac{1}{n!} f^{(n)}(a) \qquad n = 0, 1,$$

this representation is valid in the largest open disk with center a contained in D . The remainders $R_n(z)$ of (9) can be represented in the form (5). The coefficients

²COLIN MACLAURIN (1698–1746), Scots mathematician.

satisfy the inequality

$$(10) \quad |b_n| \leq \frac{M}{r^n}$$

where M is the maximum of $|f(z)|$ on the circle $|z - a| = r$.

Relation (10) follows from Cauchy's inequality (4) in Sec. 12.6.

Practically speaking, (8) means that for all z for which (9) converges, the n th partial sum of (9) will approximate $f(z)$ to any assigned degree of accuracy; we just have to choose n large enough.

From Taylor's theorem we see that the radius of convergence of (9) is at least equal to the shortest distance from a to the boundary of D . It may be larger, but then the series may no longer represent $f(z)$ at all points of D which lie in the interior of the circle of convergence.

One surprising property of complex analytic functions is that they have derivatives of all orders, and now we have discovered the other surprising property that they can always be represented by power series of the form (9). This is not true in general for real functions; there are real functions which have derivatives of all orders but cannot be represented by a power series.

The relation between our present consideration and that in Sec. 13.6 about power series may be established by the following theorem.

Theorem 2. *Every power series with a nonzero radius of convergence is the Taylor series of the function represented by that power series.*

Proof. Let the power series

$$\sum_{n=0}^{\infty} b_n(z - a)^n$$

have a nonzero radius of convergence R . Then it represents some analytic function $f(z)$ in the disk $|z - a| < R$, that is,

$$f(z) = b_0 + b_1(z - a) + b_2(z - a)^2 + \cdots.$$

From Theorem 5 in Sec. 13.6 it follows that

$$f'(z) = b_1 + 2b_2(z - a) + \cdots$$

and more generally

$$f^{(n)}(z) = n! b_n + (n + 1)n \cdots 3 \cdot 2b_{n+1}(z - a) + \cdots;$$

all these series converge in the disk $|z - a| < R$. By setting $z = a$ we obtain the following representations for the coefficients of our power series:

$$f(a) = b_0, \quad f'(a) = b_1, \quad \cdots, \quad f^{(n)}(a) = n! b_n, \quad \cdots.$$

Since these formulas are identical with those in Taylor's theorem, the proof is complete.

A point at which a function $f(z)$ ceases to be analytic is called a **singular point** of $f(z)$; we also say that $f(z)$ has a **singularity** at such a point. More precisely: a point $z = z_0$ is called a **singular point** of $f(z)$, if $f(z)$ is not differentiable at z_0 but if every neighborhood of z_0 contains points at which $f(z)$ is differentiable.

Using this concept, we may say that there is at least one singular point of $f(z)$ on the circle of convergence³ of the development (9).

14.2 Taylor Series of Elementary Functions

Example 1 (Geometric series). Let $f(z) = 1/(1 - z)$. Then we have $f^{(n)}(z) = n!/(1 - z)^{n+1}$, $f^{(n)}(0) = n!$. Hence, the Maclaurin expansion of $1/(1 - z)$ is the geometric series

$$(1) \quad \frac{1}{1 - z} = \sum_{n=0}^{\infty} z^n = 1 + z + z^2 + \dots \quad (|z| < 1).$$

$f(z)$ is singular at $z = 1$; this point lies on the circle of convergence.

Example 2 (Exponential function). We know that the exponential function e^x (Sec. 10.5) is analytic for all x , and $(e^x)' = e^x$. Hence from (9), Sec. 14.1, with $a = 0$ we obtain the Maclaurin series

$$(2) \quad e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \dots$$

This series is also obtained if we replace x in the Maclaurin series of e^x by z . We say that in this way we *continue* the real exponential function *analytically* into the complex domain.

Let us show how we can prove the multiplication formula

$$(3) \quad e^{z_1} e^{z_2} = e^{(z_1 + z_2)}$$

by using (2). We have

$$e^{z_1} e^{z_2} = \sum_{k=0}^{\infty} \frac{z_1^k}{k!} \sum_{m=0}^{\infty} \frac{z_2^m}{m!}.$$

Since both series converge absolutely, we may multiply them term by term; the sum of the products for which $k + m = n$ is

$$\begin{aligned} & \frac{z_1^n}{n!} + \frac{z_1^{n-1}}{(n-1)!} \frac{z_2}{1!} + \dots + \frac{z_1}{1!} \frac{z_2^{n-1}}{(n-1)!} + \frac{z_2^n}{n!} \\ &= \frac{1}{n!} \left[z_1^n + \binom{n}{1} z_1^{n-1} z_2 + \binom{n}{2} z_1^{n-2} z_2^2 + \dots + z_2^n \right] = \frac{(z_1 + z_2)^n}{n!}. \end{aligned}$$

Hence the product of the two series may be written

$$\sum_{n=0}^{\infty} \frac{(z_1 + z_2)^n}{n!} = e^{z_1 + z_2},$$

and (3) is proved.

³ The radius of convergence of (9) will in general be equal to the distance from a to the nearest singular point of $f(z)$, but it may be larger; for example, $\ln z$ is singular along the negative real axis, and the distance from $a = -1 + i$ to that axis is 1, but the Taylor series of $\ln z$ with $a = -1 + i$ has radius of convergence $\sqrt{2}$.

Furthermore, by setting $z = iy$ in (2) and applying Theorem 1 in Sec. 13.4 we obtain

$$e^{iy} = \sum_{n=0}^{\infty} \frac{(iy)^n}{n!} = \sum_{k=0}^{\infty} (-1)^k \frac{y^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{y^{2k+1}}{(2k+1)!}.$$

Since the series on the right are the familiar Maclaurin expansions of the real functions $\cos y$ and $\sin y$, this represents the Euler formula

$$(4) \quad e^{iy} = \cos y + i \sin y;$$

cf. (8) in Sec. 10.5. Multiplying by e^x and using (3), we obtain the formula (4) in Sec. 10.5 which was used to define e^z . Our present consideration shows that one may also use (2) for defining e^z and derive from (2) all the formulas in Sec. 10.5.

Example 3 (Trigonometric and hyperbolic functions). By substituting (2) into (1) of Sec. 10.6 we obtain

$$(5) \quad \begin{aligned} \cos z &= \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!} = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - + \cdots \\ \sin z &= \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - + \cdots. \end{aligned}$$

When $z = x$ these are the familiar Maclaurin series of the real functions $\cos x$ and $\sin x$. Similarly, by substituting (2) into (11), Sec. 10.6, we obtain

$$(6) \quad \begin{aligned} \cosh z &= \sum_{n=0}^{\infty} \frac{z^{2n}}{(2n)!} = 1 + \frac{z^2}{2!} + \frac{z^4}{4!} + \cdots \\ \sinh z &= \sum_{n=0}^{\infty} \frac{z^{2n+1}}{(2n+1)!} = z + \frac{z^3}{3!} + \frac{z^5}{5!} + \cdots. \end{aligned}$$

Example 4 (Logarithm). From (9) in Sec. 14.1 it follows that

$$(7) \quad \ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - + \cdots \quad (|z| < 1).$$

Replacing z by $-z$ and multiplying both sides by -1 , we get

$$(8) \quad -\ln(1-z) = \ln \frac{1}{1-z} = z + \frac{z^2}{2} + \frac{z^3}{3} + \cdots \quad (|z| < 1).$$

By adding both series we obtain

$$(9) \quad \ln \frac{1+z}{1-z} = 2 \left(z + \frac{z^3}{3} + \frac{z^5}{5} + \cdots \right) \quad (|z| < 1).$$

Problems

1. Differentiating (2), show that $(e^z)' = e^z$.
2. Derive (5) from (2).
3. Derive (6) from (2).
4. Derive (5) and (6) directly from (9) in Sec. 14.1.
5. Using (5) and (6), show that $\cos z$ and $\cosh z$ are even functions.
6. Derive (7) from (9) in Sec. 14.1.

Find the Taylor series of the following functions about the point $z = a$ and determine the radius of convergence.

7. $\cos 2z, a = 0$

8. $\sin z^2, a = 0$

9. $e^{-z}, a = 0$

10. $e^z, a = 1$

11. $e^z, a = \pi i$

12. $\sin z, a = \pi/2$

13. $\cos z, a = -\pi/4$

14. $1/(1-z), a = -1$

15. $1/z, a = -1$

16. $1/(1-z), a = i$

17. $\cos^2 z, a = 0$

18. $\sin^2 z, a = 0$

Find the first three terms of the Maclaurin series of the following functions.

19. $\tan z$

20. $\tan 2z$

21. $e^z \sin z$

22. $z \cot z$

Find the Maclaurin series by integrating that of the integrand term by term.

23. $\int_0^z e^t dt$

24. $\int_0^z \cos t dt$

25. $\int_0^z \sin t dt$

26. $\int_0^z \frac{e^t - 1}{t} dt$

27. $\int_0^z e^{t^2} dt$

28. $\int_0^z \frac{\sin t}{t} dt$

29. $\int_0^z \cos t^2 dt$

30. $\int_0^z \sin t^2 dt$

(The integrals in Probs. 26–30 will be considered in detail in Chap. 17.)

14.3 Practical Methods for Obtaining Power Series

In most practical cases the determination of the coefficients of a Taylor series by means of the formula in Taylor's theorem will be complicated or time-consuming. There are a number of simpler practical procedures for that purpose which may be illustrated by the following examples. The uniqueness of the representations thus obtained follows from Theorem 2 in Sec. 13.6.

Example 1 (Substitution). Find the Maclaurin series of $f(z) = 1/(1 + z^2)$. By substituting $-z^2$ for z in (1), Sec. 14.2, we obtain

$$(1) \quad \frac{1}{1+z^2} = \frac{1}{1-(-z^2)} = \sum_{n=0}^{\infty} (-z^2)^n = \sum_{n=0}^{\infty} (-1)^n z^{2n} \\ = 1 - z^2 + z^4 - z^6 + \dots \quad (|z| < 1).$$

Example 2 (Integration). Let $f(z) = \tan^{-1} z$. We have $f'(z) = 1/(1 + z^2)$. By integrating (1) term by term and noting that $f(0) = 0$ we find

$$\tan^{-1} z = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} z^{2n+1} = z - \frac{z^3}{3} + \frac{z^5}{5} - \dots \quad (|z| < 1);$$

this series represents the principal value of $w = u + iv = \tan^{-1} z$, defined as that value for which $|u| < \pi/2$.

Example 3 (Development by using the geometric series). Develop $1/(c - bz)$ in powers of $z - a$ where $c - ab \neq 0$ and $b \neq 0$. Obviously,

$$\frac{1}{c - bz} = \frac{1}{c - ab - b(z - a)} - \frac{1}{(c - ab) \left[1 - \frac{b(z - a)}{c - ab} \right]}.$$

To the last expression we apply (1) in Sec. 14.2 with z replaced by $b(z - a)/(c - ab)$, finding

$$\frac{1}{c - bz} = \frac{1}{c - ab} \sum_{n=0}^{\infty} \left[\frac{b(z - a)}{(c - ab)} \right]^n = \sum_{n=0}^{\infty} \frac{b^n}{(c - ab)^{n+1}} (z - a)^n.$$

Writing the last series at length, we have

$$\frac{1}{c - bz} = \frac{1}{c - ab} + \frac{b}{(c - ab)^2} (z - a) + \frac{b^2}{(c - ab)^3} (z - a)^2 + \dots$$

This series converges for

$$\left| \frac{b(z - a)}{c - ab} \right| < 1 \quad \text{that is,} \quad |z - a| < \left| \frac{c - ab}{b} \right| = \frac{c}{b} - a$$

Example 4 (Binomial series, reduction by partial fractions). Find the Taylor series of the function

$$f(z) = \frac{2z^3 + 9z + 5}{z^3 + z^2 - 8z - 12}$$

with center at $z = 1$.

Given a rational function, we may first represent it as a sum of partial fractions and then apply the *binomial series*

$$(2) \quad \begin{aligned} \frac{1}{(1+z)^m} &= (1+z)^{-m} = \sum_{n=0}^{\infty} \binom{-m}{n} z^n \\ &= 1 - mz + \frac{m(m+1)}{2!} z^2 - \frac{m(m+1)(m+2)}{3!} z^3 + \dots \end{aligned}$$

Since the function on the left is singular at $z = -1$, the series converges in the disk $|z| < 1$. In our case we obtain

$$f(z) = \frac{1}{(z+2)^2} + \frac{2}{z-3} = \frac{1}{[3+(z-1)]^2} - \frac{2}{2-(z-1)}.$$

This may be written in the form

$$f(z) = \frac{1}{9} \left(\frac{1}{[1+\frac{1}{3}(z-1)]^2} - \frac{1}{1-\frac{1}{2}(z-1)} \right).$$

By using the binomial series we obtain

$$f(z) = \frac{1}{9} \sum_{n=0}^{\infty} \binom{-2}{n} \left(\frac{z-1}{3} \right)^n - \sum_{n=0}^{\infty} \binom{z-1}{2}^n.$$

We may add the two series on the right term by term. Since the binomial coefficient in the first series equals $(-1)^n(n+1)$, we find

$$f(z) = \sum_{n=0}^{\infty} \left[\frac{(-1)^n(n+1)}{3^{n+2}} - \frac{1}{2^n} \right] (z-1)^n.$$

Writing this result at length, we have

$$f(z) = -\frac{8}{9} - \frac{31}{54}(z-1) - \frac{83}{108}(z-1)^2 - \dots$$

Since $z = 3$ is the singular point of $f(z)$ which is nearest to the center $z = 1$, the series converges in the disk $|z-1| < 2$.

Example 5 (Use of differential equations). Find the Maclaurin series of $f(z) = \tan z$. We have $f'(z) = \sec^2 z$ and, therefore,

$$f'(z) = 1 + f^2(z), \quad f'(0) = 1.$$

Observing that $f(0) = 0$, we obtain by successive differentiation

$$\begin{aligned} f'' &= 2ff', & f''(0) &= 0, \\ f''' &= 2f'^2 + 2ff'', & f'''(0) &= 2, & f'''(0)/3! &= 1/3, \\ f^{(4)} &= 6f'f'' + 2f'''^2, & f^{(4)}(0) &= 0, \\ f^{(5)} &= 6f'^2 + 8ff''' + 2f^{(4)}, & f^{(5)}(0) &= 16, & f^{(5)}(0)/5! &= 2/15, & \text{etc.} \end{aligned}$$

Hence the result is

$$(3) \quad \tan z = z + \frac{1}{3}z^3 + \frac{2}{15}z^5 + \frac{17}{315}z^7 + \dots \quad \left(|z| < \frac{\pi}{2}\right).$$

Example 6 (Undetermined coefficients). Find the Maclaurin series of $\tan z$ by using those of $\cos z$ and $\sin z$ (Sec. 14.2). Since $\tan z$ is odd, the desired expansion will be of the form

$$\tan z = b_1 z + b_3 z^3 + b_5 z^5 + \dots$$

Using $\sin z = \tan z \cos z$ and inserting those developments, we obtain

$$z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots = (b_1 z + b_3 z^3 + b_5 z^5 + \dots) \left(1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots\right).$$

Since $\tan z$ is analytic except at $z = \pm\pi/2, \pm 3\pi/2, \dots$, its Maclaurin series converges in the disk $|z| < \pi/2$, and for these z we may multiply the two series on the right term by term and arrange the resulting series in powers of z (cf. Theorem 3 in Sec. 13.5). By Theorem 2 in Sec. 13.6 the coefficient of each power of z is the same on both sides. This yields

$$1 = b_1, \quad -\frac{1}{3!} = -\frac{b_1}{2!} + b_3, \quad \frac{1}{5!} = \frac{b_1}{4!} - \frac{b_3}{2!} + b_5, \quad \text{etc.}$$

Therefore $b_1 = 1$, $b_3 = \frac{1}{3}$, $b_5 = \frac{2}{15}$, etc., as before.

We mention that there exist tables⁴ of the so-called Bernoulli numbers B_n from which the coefficients of (3) can easily be calculated. The numbers $B_n/n!$ are, by definition, the coefficients of the Maclaurin series

$$(4) \quad \frac{z}{e^z - 1} = 1 + B_1 z + \frac{B_2}{2!} z^2 + \frac{B_3}{3!} z^3 + \dots$$

By the method of undetermined coefficients we obtain

$$(5) \quad B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_3 = 0, \quad B_4 = -\frac{1}{30}, \quad B_5 = 0, \quad B_6 = \frac{1}{42}, \dots$$

From (1), (2), Sec. 10.6, it follows that

$$\tan z = \frac{2i}{e^{2iz} - 1} - \frac{4i}{e^{4iz} - 1} - i,$$

as the student may verify. From this and (4) we obtain

$$) \quad \tan z = \frac{4 \cdot 3}{2!} B_2 z + \dots + (-1)^{n-1} \frac{2^{2n}(2^{2n}-1)}{(2n)!} B_{2n} z^{2n-1} + \dots$$

⁴ Cf. J. W. L. Glaisher, Quarterly Journal Math. 29, 1898, 1-168.

Problems

1. Obtain the series in Ex. 2 from Prob. 28 in Sec. 10.7 and (9) in Sec. 14.2.

Find the Maclaurin series of the following functions.

2. $\frac{1}{1-z^3}$

3. $\frac{1}{1+z^3}$

4. $\frac{1}{1-z^6}$

6. $\sin z^3$

7. $\cos z^2$

8. e^{z^3-z}

5. $\frac{1}{(1+z^2)^2}$

9. e^{z^4}

10. $\frac{\sqrt{z}}{2} \int_0^z \frac{\cos t}{\sqrt{t}} dt$

11. $\frac{\sqrt{z}}{2} \int_0^z \frac{\sin t}{\sqrt{t}} dt$

12. $e^{z^2} \int_0^z e^{-t^2} dt$

13. Developing $1/\sqrt{1-z^2}$ and integrating, show that

$$\sin^{-1} z = z + \left(\frac{1}{2}\right)\frac{z^3}{3} + \left(\frac{1 \cdot 3}{2 \cdot 4}\right)\frac{z^5}{5} + \left(\frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\right)\frac{z^7}{7} + \dots \quad (|z| < 1).$$

14. Show that the series in Prob. 13 represents the principal value of $\sin^{-1} z$ (cf. Prob. 30 in Sec. 10.7).

Find the first few terms of the Maclaurin series of the following functions.

15. $e^{(e^z)}$

16. $\cos \left(\frac{z}{1-z} \right)$

17. $e^{1/(1-z)}$

Find the Taylor series expansion of the given function about $z = a$.

18. $\frac{1}{2z-i}$, $a = -1$

19. $\frac{1}{4-3z}$, $a = 1+i$

20. $\frac{1}{1-z}$, $a = 2i$

21. $\frac{1}{(1+z)^2}$, $a = -i$

22. $\frac{1}{(2+3z^3)^2}$, $a = 0$

23. $\tan z$, $a = \frac{\pi}{4}$

24. Derive (6) from (4).

25. (Euler numbers) The expansion

$$\sec z = E_0 - \frac{E_2}{2!} z^2 + \frac{E_4}{4!} z^4 - + \dots$$

defines the Euler numbers E_{2n} . Show that $E_0 = 1$, $E_2 = -1$, $E_4 = 5$, $E_6 = -61$.

14.4 Uniform Convergence

Suppose we know that a given series converges in a certain region R . Then the remaining question is whether the convergence is sufficiently rapid throughout the whole region or whether there are points near which the convergence becomes poor. The practical importance of this question in connection with computations is obvious, but we shall see that the theoretical aspect of the question is even more important. To illustrate the situation, let us start with some examples.

Example 1. Suppose we want to compute a table of e^x for real x in the interval $0 \leq x \leq 1$, for example, for $x = 0, 0.1, 0.2, \dots$ and the absolute value of the error of each value should be less than a given number ϵ , say, less than $\frac{1}{2}$ unit of the sixth decimal place. We may use a suitable partial sum

$$s_n = 1 + x + \cdots + \frac{x^n}{n!}$$

of the MacLaurin series. Then the absolute value of the error is equal to $|R_n| = |s - s_n|$ where $s = e^x$, the sum of the series, and we have to choose n such that

$$|s(x) - s_n(x)| < \epsilon \quad (= 5 \cdot 10^{-7}).$$

From Prob. 20 in Sec. 13.3 it follows that when $x = 1$, then for $n = 10$, and thus for any $n > N = 9$, we obtain the required accuracy. Now the remainder decreases in absolute value as $x (\geq 0)$ decreases, and, therefore,

$$|s(x) - s_n(x)| < \epsilon \quad \text{for} \quad n > N(\epsilon) (= 9) \quad \text{and all } x$$

under consideration. Note that, of course, N depends on ϵ , and if we want more accurate values so that ϵ is smaller, then N will be larger.

Example 2. In the case of the geometric series $1 + z + z^2 + \cdots$ the remainder is

$$R_n(z) = s(z) - s_n(z) = \sum_{m=n+1}^{\infty} z^m = \frac{z^{n+1}}{1-z}$$

and becomes arbitrarily large for real $z = x (< 1)$ and sufficiently close to 1. Hence, a maximum error ϵ being prescribed, we cannot find an N depending only on ϵ such that $|R_n(x)| = |s(x) - s_n(x)| < \epsilon$ for $n > N$ and all x in the interval $0 \leq x < 1$. Cf. also Prob. 21 at the end of Sec. 13.3.

The result in Ex. 2 is not quite unexpected, because the series diverges at $z = 1$. A really surprising situation arises in the case of the following series.

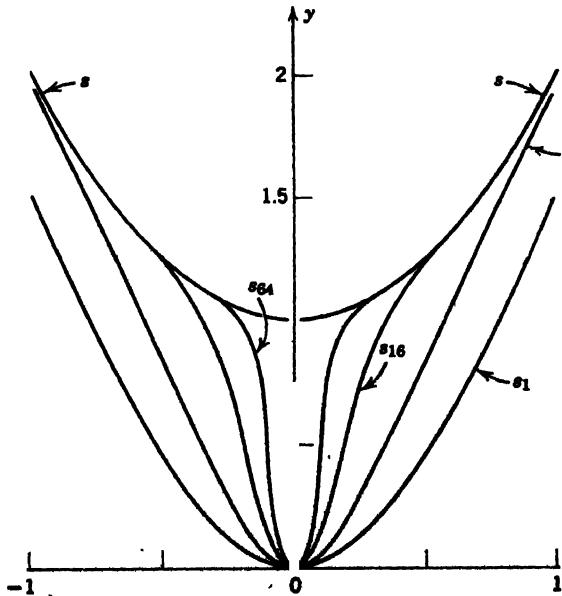


Fig. 343. Partial sums and sum in Example 3.

Example 3. Consider the series

$$x^3 + \frac{x^3}{1+x^3} + \frac{x^3}{(1+x^3)^2} + \frac{x^3}{(1+x^3)^3} + \dots$$

Using the formula for the sum of a geometric progression, the student may readily verify that the n th partial sum is

$$s_n(x) = 1 + x^3 - \frac{1}{(1+x^3)^n}.$$

Hence, if $x \neq 0$, the series has the sum

$$s(x) = \lim_{n \rightarrow \infty} s_n(x) = 1 + x^3.$$

If $x = 0$, then $s_n = 0$ for all n and, therefore,

$$s(0) = \lim_{n \rightarrow \infty} s_n(0) = 0.$$

This shows that the series converges for all x (even absolutely), but we have the surprising result that the sum is discontinuous (at $x = 0$), although all the terms of the series are continuous functions. Furthermore, when $x \neq 0$ the absolute value of the remainder is

$$|R_n(x)| = |s(x) - s_n(x)| = \frac{1}{(1+x^3)^n}$$

and we see that for a given $\epsilon (< 1)$ we cannot find an N depending only on ϵ such that $|R_n| < \epsilon$ for all $n > N(\epsilon)$ and all x in the interval $0 \leq x \leq 1$.

The series in the examples are of the form

$$(1) \quad \sum_{n=0}^{\infty} f_n(z) = f_0(z) + f_1(z) + f_2(z) + \dots$$

We assume that (1) converges for all z in a region R . Let $s(z)$ be the sum and $s_n(z)$ the n th partial sum of (1). We know that convergence of (1) at a point z means that, an $\epsilon > 0$ being given, we can find an $N = N(\epsilon, z)$ such that

$$|s(z) - s_n(z)| < \epsilon \quad \text{for all } n > N(\epsilon, z).$$

N depends on ϵ and will, in general, also depend on the point z which has been selected for consideration. Now, any $\epsilon > 0$ being given, it may happen that we can find a number $N(\epsilon)$, independent of z , such that

$$|s(z) - s_n(z)| < \epsilon \quad \text{for all } n > N(\epsilon) \text{ and all } z \text{ in } R.$$

Then the series is said to be uniformly convergent in R .

Uniformity of convergence is thus a property that depends on a whole set of values of z , whereas the convergence of a series may be considered for various particular values of z , without reference to other values.

The series in Ex. 1 is uniformly convergent in the interval $0 \leq x \leq 1$ (and, in fact, in any bounded region of the z -plane), while the series in Ex. 3 is not uniformly convergent in a region containing the point 0. This shows

that an absolutely convergent series may not be uniformly convergent. Conversely, uniformly convergent series may not be absolutely convergent. This is illustrated by

Example 4. The series

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{z^2 + n} = \frac{1}{z^2 + 1} - \frac{1}{z^2 + 2} + \frac{1}{z^2 + 3} - + \cdots \quad (z \text{ real})$$

is uniformly convergent for all real z , but not absolutely convergent (cf. Prob. 6).

Example 2 is typical for power series, because for such a series the situation is very simple, as follows.

Theorem 1. A power series

$$(2) \quad \sum_{n=0}^{\infty} c_n(z - a)^n$$

with a nonzero radius of convergence R is uniformly convergent in every circular disk $|z - a| \leq r$ of radius $r < R$.

Proof. For $|z - a| \leq r$ we have

$$(3) \quad \begin{aligned} & |c_{n+1}(z - a)^{n+1} + \cdots + c_{n+p}(z - a)^{n+p}| \\ & \leq |c_{n+1}| r^{n+1} + \cdots + |c_{n+p}| r^{n+p}. \end{aligned}$$

Since (2) converges absolutely at $z = r$, it follows from Theorem 4 in Sec. 13.2 that, an $\epsilon > 0$ being given, we can find an $N(\epsilon)$ such that

$$|c_{n+1}| r^{n+1} + \cdots + |c_{n+p}| r^{n+p} < \epsilon \quad \text{for } n > N(\epsilon) \quad \text{and } p = 1, 2, \dots$$

From this and (3) we obtain

$$|c_{n+1}(z - a)^{n+1} + \cdots + c_{n+p}(z - a)^{n+p}| < \epsilon$$

for all z in the disk $|z - a| \leq r$, every $n > N(\epsilon)$, and every $p = 1, 2, \dots$. Since $N(\epsilon)$ is independent of z , this shows uniform convergence, and the theorem is proved.

While, of course, the sum of finitely many continuous functions is continuous, Ex. 3 illustrates that the sum of an infinite series of continuous functions may be discontinuous, even if it converges absolutely. But if a series converges uniformly, this cannot happen. In fact the following important theorem holds.

Theorem 2. Let the series

$$\sum_{m=0}^{\infty} f_m(z) = f_0(z) + f_1(z) + \cdots$$

be uniformly convergent in a region R and let $F(z)$ be its sum. Then, if each term $f_m(z)$ is continuous at a point z_0 in R , the function $F(z)$ is continuous at z_0 .

Proof. Let $s_n(z)$ be the n th partial sum of the series and $R_n(z)$ the corresponding remainder:

$$s_n = f_0 + f_1 + \cdots + f_n, \quad R_n = f_{n+1} + f_{n+2} + \cdots.$$

An $\epsilon > 0$ being given, we can find an $n = N(\epsilon)$ such that

$$|R_N(z)| < \frac{\epsilon}{3} \quad \text{for all } z \text{ in } R,$$

because the series converges uniformly. Since $s_N(z)$ is a sum of finitely many functions which are continuous at z_0 , this sum is continuous at z_0 . Therefore we can find a $\delta > 0$ such that

$$|s_N(z) - s_N(z_0)| < \frac{\epsilon}{3} \quad \text{for all } z \text{ in } R \text{ for which } |z - z_0| < \delta.$$

By the triangle inequality (Sec. 10.1) for these z we thus obtain

$$\begin{aligned} |F(z) - F(z_0)| &= |s_N(z) + R_N(z) - [s_N(z_0) + R_N(z_0)]| \\ &\leq |s_N(z) - s_N(z_0)| + |R_N(z)| + |R_N(z_0)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \end{aligned}$$

This implies that $F(z)$ is continuous at z_0 , and the theorem is proved.

We want to mention that in this theorem uniformity of convergence is a sufficient rather than a necessary condition. This may be illustrated by the following example.

Example 5. Let

$$u_m(x) = \frac{mx}{1 + m^2x^2}$$

and consider the series

$$\sum_{m=1}^{\infty} f_m(x) \quad \text{where} \quad f_m(x) = u_m(x) - u_{m-1}(x).$$

The n th partial sum is

$$s_n = u_1 - u_0 + u_2 - u_1 + \cdots + u_n - u_{n-1} = u_n - u_0 = u_n.$$

Hence the series has the sum

$$F(x) = \lim_{n \rightarrow \infty} s_n(x) = \lim_{n \rightarrow \infty} u_n(x) = 0,$$

which is a continuous function. However, the series is not uniformly convergent in an interval $0 \leq x \leq a$, where $a > 0$. In fact, from

$$|F(x) - s_n(x)| = \frac{nx}{1 + n^2x^2} < \epsilon$$

we obtain

$$\frac{nx}{\epsilon} < 1 + n^2x^2 \quad \text{or} \quad n^2x^2 - \frac{nx}{\epsilon} + 1 > 0$$

and from this

$$n > \frac{1}{2x\epsilon} (1 + \sqrt{1 - 4\epsilon^2}).$$

For fixed ϵ the right side approaches infinity as x approaches zero, which shows that the series is not uniformly convergent in that interval.

Under what conditions may we integrate a series term by term?

Let us start our consideration with an example which illustrates the important fact that term by term integration of series is not always permissible.

Example 6. Let

$$u_m(x) = mx e^{-mx^2}$$

and consider the series

$$\sum_{m=1}^{\infty} f_m(x) \quad \text{where} \quad f_m(x) = u_m(x) - u_{m-1}(x)$$

in the interval $0 \leq x \leq 1$. The n th partial sum is

$$s_n = u_1 - u_0 + u_2 - u_1 + \cdots + u_n - u_{n-1} = u_n - u_0 = u_n.$$

Hence the series has the sum

$$F(x) = \lim_{n \rightarrow \infty} s_n(x) = \lim_{n \rightarrow \infty} u_n(x) = 0 \quad (0 \leq x \leq 1).$$

From this we obtain

$$\int_0^1 F(x) dx = 0.$$

On the other hand, by integrating term by term,

$$\sum_{m=1}^{\infty} \int_0^1 f_m(x) dx = \lim_{n \rightarrow \infty} \sum_{m=1}^n \int_0^1 f_m(x) dx = \lim_{n \rightarrow \infty} \int_0^1 s_n(x) dx.$$

Now $s_n = u_n$ and the expression on the right becomes

$$\lim_{n \rightarrow \infty} \int_0^1 u_n(x) dx = \lim_{n \rightarrow \infty} \int_0^1 nx e^{-nx^2} dx = \lim_{n \rightarrow \infty} \frac{1}{2}(1 - e^{-n}) = \frac{1}{2},$$

but not 0. This shows that the series under consideration cannot be integrated term by term from $x = 0$ to $x = 1$.

The series in Ex. 6 is not uniformly convergent in that interval, and we shall now prove that in the case of a uniformly convergent series of continuous functions we may integrate term by term.

Theorem 3. *Let*

$$F(z) = \sum_{n=0}^{\infty} f_n(z) = f_0(z) + f_1(z) + \cdots$$

be a uniformly convergent series of continuous functions within a region R . Let C be any path in R . Then the series

$$(4) \quad \sum_{n=0}^{\infty} \int_C f_n(z) dz = \int_C f_0(z) dz + \int_C f_1(z) dz + \cdots$$

is convergent and has the sum $\int_C F(z) dz$.

Proof. From Theorem 2 it follows that $F(z)$ is continuous. Let $s_n(z)$ be the n th partial sum of the given series and $R_n(z)$ the corresponding remainder. Then $F = s_n + R_n$ and

$$\int_C F(z) dz = \int_C s_n(z) dz + \int_C R_n(z) dz.$$

Let l be the length of C . Since the given series converges uniformly, to every given $\epsilon > 0$ we can find a number N such that

$$|R_n(z)| < \frac{\epsilon}{l} \quad \text{for all } n > N \text{ and all } z \text{ in } R.$$

By applying (4), Sec. 12.2, we thus obtain

$$\left| \int_C R_n(z) dz \right| < \frac{\epsilon}{l} l = \epsilon \quad \text{for all } n > N.$$

Since $R_n = F - s_n$, this means that

$$\left| \int_C F(z) dz - \int_C s_n(z) dz \right| < \epsilon \quad \text{for all } n > N.$$

Hence, the series (4) converges and has the sum indicated in the theorem. This completes the proof.

Theorem 2 and Theorem 3 characterize the two most important properties of uniformly convergent series.

Of course, since differentiation and integration are inverse processes, we readily conclude from Theorem 3 that a convergent series may be differentiated term by term, provided the terms of the given series have continuous derivatives and the resulting series is uniformly convergent; more precisely, the following theorem holds.

Theorem 4. Suppose that the series $f_0(z) + f_1(z) + f_2(z) + \dots$ is convergent in a region R and has the sum $F(z)$, the derivatives $f_n'(z)$ are continuous in R , and the series $f_0'(z) + f_1'(z) + f_2'(z) + \dots$ converges uniformly in R . Then

$$F'(z) = f_0'(z) + f_1'(z) + f_2'(z) + \dots \quad \text{for all } z \text{ in } R.$$

The simple proof is left to the student (Prob. 8).

Usually uniform convergence is established by a comparison test, namely, by the so-called

Weierstrass M-test. If, for all values of z within a region R , the absolute values of the terms of a given series of the form (1) are, respectively, less than the corresponding terms in a convergent series of constant terms,

$$(5) \quad M_0 + M_1 + M_2 + \dots,$$

then the series (1) converges uniformly in R .

The simple proof is left to the student (Prob. 7).

Example 7. The Weierstrass test shows that the series

$$\sum_{m=1}^{\infty} \frac{\sin mx}{m^2} \quad (x \text{ real})$$

converges uniformly in any interval, because for real x ,

$$\left| \frac{\sin mx}{m^2} \right| \leq \frac{1}{m^2}$$

and $\sum m^{-2}$ is convergent [cf. (5) in Sec. 13.3].

Differentiating the series term by term, we obtain

$$\sum_{m=1}^{\infty} \frac{\cos mx}{m} .$$

At $x = 0$ this series becomes the harmonic series which is not convergent. This shows that the last assumption in Theorem 4 cannot be omitted.

Problems

1. Determine the smallest integer n such that $|R_n| < 0.01$ in Ex. 2, when $x = 0.5, 0.6, 0.7, 0.8, 0.9$. What does the result mean from the viewpoint of computing $1/(1-x)$ with an absolute error less than 0.01 by means of the geometric series?
2. Graph s_1, s_2, \dots, s_5 , and s in Ex. 2 for $-1 < x < 1$.
3. In Ex. 2, show that $R_{2n}(x) \rightarrow -1/2$ and $R_{2n+1}(x) \rightarrow 1/2$ as $x \rightarrow -1$. Conclude from the result that the convergence of the geometric series is not uniform in the interval $-1 < x < 0$.
4. Prove that the series in Ex. 3 is not uniformly convergent in any interval containing the point $x = 0$.
5. Show that if the series (1) is uniformly convergent in some region R , it is uniformly convergent in any portion of R .
6. Prove the statement in Ex. 4.
7. Give a proof of the Weierstrass M-test
8. Derive Theorem 4 from Theorem 3.
9. Show that $1 + \sum_{n=1}^{\infty} (x^n - x^{n-1})$ is not uniformly convergent in the interval $0 \leq x \leq 1$. Graph the partial sums s_1, s_2, s_3, s_4 .

10. Show that $\sum_{n=1}^{\infty} \left(\frac{x^{2n}}{1+x^{2n}} - \frac{x^{2n-2}}{1+x^{2n-2}} \right) = \begin{cases} -1 & \text{when } |x| < 1 \\ -\frac{1}{2} & \text{when } x = \pm 1 \\ 0 & \text{when } |x| > 1 \end{cases}$

Hint: consider the partial sums.

Prove that the following series converge uniformly in the given regions R . (Here x is real.)

11. $\sum_{n=0}^{\infty} x^n, \quad R: |x| \leq 0.9$

12. $\sum_{n=1}^{\infty} \frac{\cos nx}{2^n}, \quad R: \text{all } x$

13. $\sum_{n=0}^{\infty} \frac{\tanh^n x}{n!}, \quad R: \text{all } x$

14. $\sum_{n=1}^{\infty} \frac{\cos^n x}{n^2}, \quad R: \text{all } x$

$$15. \sum_{n=1}^{\infty} \frac{1}{|z| + n^2}, \quad R: \text{all } z$$

$$17. \sum_{n=1}^{\infty} \frac{z^n}{n^2}, \quad R: |z| \leq 1$$

$$16. \sum_{n=1}^{\infty} \frac{\sin n|z|}{n(n+1)}, \quad R: \text{all } z$$

$$18. \sum_{n=0}^{\infty} \frac{z^n}{n!}, \quad R: |z| \leq 10^{50}$$

Show that (10), Sec. 9.5, with coefficients (11) is a solution of the heat equation for $t > 0$, assuming that $f(x)$ is continuous on the interval $0 \leq x \leq l$ and has one-sided derivatives at all interior points of that interval. Proceed as follows.

19. Show that $|B_n|$ is bounded, say, $|B_n| < K$ for all n . Conclude that

$$|u_n| < Ke^{-\lambda_n^2 t_0} \quad \text{when } t \geq t_0 > 0$$

and, by the Weierstrass test, the series (10) converges uniformly with respect to x and t when $t \geq t_0$, $0 \leq x \leq l$. Using Theorem 2, show that $u(x, t)$ is continuous when $t \geq t_0$ and thus satisfies the boundary conditions (2) when $t \geq t_0$.

20. Show that $|\partial u_n / \partial t| < \lambda_n^2 Ke^{-\lambda_n^2 t_0}$ when $t \geq t_0$ and the series of the expressions on the right converges, by the ratio test. Conclude from this, the Weierstrass test, and Theorem 4 that the series (10) can be differentiated term by term with respect to t and the resulting series has the sum $\partial u / \partial t$. Show that (10) can be differentiated twice with respect to x and the resulting series has the sum $\partial^2 u / \partial x^2$. Conclude from this and the result of Prob. 19 that (10) is a solution of the heat equation for all $t \geq t_0$. (The proof that (10) satisfies the given initial condition can be found in Ref. [E2].)

14.5 Laurent Series

In various applications it is necessary to expand a function $f(z)$ around points where $f(z)$ is singular. Taylor's theorem cannot be applied in such cases. A new type of series, known as *Laurent series*, is necessary. This will be a representation which is valid in an annulus bounded by two concentric circles C_1 and C_2 such that $f(z)$ is analytic in the annulus and at each point of C_1 and C_2 (Fig. 344). As in the case of the Taylor series, $f(z)$ may be singular at some points outside C_1 and, as the essentially new feature, it may also be singular at some points inside C_2 .

Laurent's theorem.⁵ *If $f(z)$ is analytic and single-valued on two concentric circles C_1 and C_2 with center a and in the annulus between them, then $f(z)$ can be represented by the Laurent series*

$$(1) \quad f(z) = \sum_{n=0}^{\infty} b_n(z-a)^n + \sum_{n=1}^{\infty} \frac{c_n}{(z-a)^n}$$

$$= b_0 + b_1(z-a) + b_2(z-a)^2 + \cdots + \frac{c_1}{z-a} + \frac{c_3}{(z-a)^3} + \cdots$$

where⁶

$$(2) \quad b_n = \frac{1}{2\pi i} \int_C \frac{f(z^*)}{(z^* - a)^{n+1}} dz^*, \quad c_n = \frac{1}{2\pi i} \int_C (z^* - a)^{n-1} f(z^*) dz^*,$$

⁵ HERMANN LAURENT (1841–1908), French mathematician.

⁶ We denote the variable of integration by z^* because z is used in $f(z)$.

each integral being taken counterclockwise around any simple closed path C which lies in the annulus and encircles the inner circle (Fig. 344).

This series converges and represents $f(z)$ in the open annulus obtained from the given annulus by continuously increasing the circle C_1 and decreasing C_2 until each of the two circles reaches a point where $f(z)$ is singular.

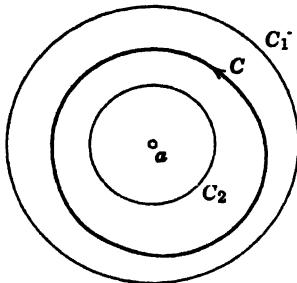


Fig. 344. Laurent's theorem.

Remark. Obviously, instead of (1) and (2) we may write simply

$$(1') \quad f(z) = \sum_{n=-\infty}^{\infty} A_n(z-a)^n$$

where

$$(2') \quad A_n = \frac{1}{2\pi i} \int_C \frac{f(z^*)}{(z^*-a)^{n+1}} dz^*.$$

Proof of Laurent's theorem. Let z be any point in the given annulus. Then from Cauchy's integral formula [cf. (3) in Sec. 12.5] it follows that

$$(3) \quad f(z) = \frac{1}{2\pi i} \int_{C_1} \frac{f(z^*)}{z^*-z} dz^* - \frac{1}{2\pi i} \int_{C_2} \frac{f(z^*)}{z^*-z} dz^*,$$

where we integrate in the counterclockwise sense. We shall now transform these integrals in a fashion similar to that in Sec. 14.1. Since z lies inside C_1 , the first of these integrals is precisely of the same type as the integral (1), Sec. 14.1. By expanding it and estimating the remainder as in Sec. 14.1 we obtain

$$(4) \quad \frac{1}{2\pi i} \int_{C_1} \frac{f(z^*)}{z^*-z} dz^* = \sum_{n=0}^{\infty} b_n(z-a)^n$$

where the coefficients are given by the formula

$$(5) \quad b_n = \frac{1}{2\pi i} \int_{C_1} \frac{f(z^*)}{(z^*-a)^{n+1}} dz^*$$

and we integrate in the counterclockwise sense. Since a is not a point of the annulus, the functions $f(z^*)/(z^*-a)^{n+1}$ are analytic in the annulus. Hence, we may integrate along the path C (cf. in the theorem) instead of C_1 , without altering the value of b_n . This proves (2) for all $n \geq 0$.

In the case of the last integral in (3), the situation is different, since z lies outside C_2 . Instead of (3), Sec. 14.1, we now have

$$z^* - a < 1,$$

that is, we now have to develop $1/(z^* - z)$ in powers of $(z^* - a)/(z - a)$ for the resulting series to be convergent. We find

$$z^* - z = \frac{1}{z^* - a - (z - a)} = \frac{-1}{(z - a) \left(1 - \frac{z^* - a}{z - a} \right)}.$$

By applying the formula for a finite geometric progression to the last expression we obtain

$$\begin{aligned} \frac{1}{z^* - z} &= -\frac{1}{z - a} \left\{ 1 + \frac{z^* - a}{z - a} + \left(\frac{z^* - a}{z - a} \right)^2 + \cdots + \left(\frac{z^* - a}{z - a} \right)^n \right\} \\ &\quad - \frac{1}{z - z^*} \left(\frac{z^* - a}{z - a} \right)^{n+1}. \end{aligned}$$

From this development we readily have

$$\begin{aligned} &-\frac{1}{2\pi i} \int_{C_2} \frac{f(z^*)}{z^* - z} dz^* \\ &= \frac{1}{2\pi i} \left\{ \frac{1}{z - a} \int_{C_1} f(z^*) dz^* + \frac{1}{(z - a)^2} \int_{C_2} (z^* - a) f(z^*) dz^* + \cdots \right. \\ &\quad \left. + \frac{1}{(z - a)^{n+1}} \int_{C_2} (z^* - a)^n f(z^*) dz^* \right\} + R_n^*(z); \end{aligned}$$

in this representation the last term is of the form

$$(7) \quad R_n^*(z) = \frac{1}{2\pi i (z - a)^{n+1}} \int_{C_2} \frac{(z^* - a)^{n+1}}{z - z^*} f(z^*) dz^*.$$

In the integrals on the right we may replace the circle C_2 by the aforementioned path C , without altering their values. This establishes Laurent's theorem provided that

$$(8) \quad \lim_{n \rightarrow \infty} R_n^*(z) = 0.$$

We prove (8). Since $z - z^* \neq 0$ and $f(z)$ is analytic in the annulus and on C_2 , the absolute value of the expression $f(z^*)/(z - z^*)$ in (7) is bounded, say,

$$\left| \frac{f(z^*)}{z - z^*} \right| < \tilde{M} \quad \text{for all } z^* \text{ on } C_2.$$

By applying (4) in Sec. 12.2 to (7) and denoting the length of C_2 by l we thus obtain

$$|R_n^*(z)| < \frac{1}{2\pi |z - a|^{n+1}} |z^* - a|^{n+1} \tilde{M} l = \frac{\tilde{M} l}{2\pi} \left| \frac{z^* - a}{z - a} \right|^{n+1}.$$

From (6) we see that the expression on the right approaches zero as n approaches infinity. This proves (8). The representation (1) with coefficients (2) is now established in the given annulus.

Finally let us prove convergence of (1) in the open annulus characterized at the end of the theorem.

We denote the sums of the two series in (1) by $g(z)$ and $h(z)$, and the radii of C_1 and C_2 by r_1 and r_2 , respectively. Then $f = g + h$. The first series is a power series. Since it converges in the annulus, it must converge in the entire disk bounded by C_1 , and g is analytic in that disk.

Setting $Z = 1/(z - a)$, the last series becomes a power series in Z . The annulus $r_2 < |z - a| < r_1$ then corresponds to the annulus $1/r_1 < |Z| < 1/r_2$, the new series converges in this annulus and, therefore, in the entire disk $|Z| < 1/r_2$. Since this disk corresponds to $|z - a| > r_2$, the exterior of C_2 , the given series converges for all z outside C_2 , and h is analytic for all these z .

Since $f = g + h$, it follows that g must be singular at all those points outside C_1 where f is singular, and h must be singular at all those points inside C_2 where f is singular. Consequently, the first series converges for all z inside the circle about a whose radius is equal to the distance of that singularity of f outside C_1 which is closest to a . Similarly, the second series converges for all z outside the circle about a whose radius is equal to the maximum distance of the singularities of f inside C_2 . The domain common to both of those domains of convergence is the open annulus characterized at the end of the theorem, and the proof is complete.

It follows that if $f(z)$ is analytic inside C_2 , the Laurent series reduces to the Taylor series of $f(z)$ with center a . In fact, by applying Cauchy's integral theorem to (2) we see that in this case all the coefficients of the negative powers in (1) are zero.

Furthermore, if $z = a$ is the only singular point of $f(z)$ in C_2 , then the Laurent expansion (1) converges for all z in C_1 except at $z = a$. This case occurs frequently and, therefore, is of particular importance.

The Laurent series of a given analytic function $f(z)$ in its annulus of convergence is unique (cf. Prob. 9, this section). *However, $f(z)$ may have different Laurent series in two annuli with the same center* (cf. Ex. 2, below).

The uniqueness is important, because Laurent series usually are not obtained by using (2) for determining the coefficients, but by various other methods. Some of these methods are illustrated by the following examples. If a Laurent series is found by any such process, it must be the Laurent series of the given function in the given annulus.

Example 1. The Laurent series of $z^2 e^{1/z}$ with center 0 can be obtained from (2) in Sec. 14.2. Replacing z by $1/z$ in that series, we find

$$z^2 e^{1/z} = z^2 \left(1 + \frac{1}{1! z} + \frac{1}{2! z^2} + \cdots \right) = z^2 + z + \frac{1}{2} + \frac{1}{3! z} + \frac{1}{4! z^2} + \cdots \quad (|z| > 0).$$

Example 2. Find all Laurent series of $f(z) = 1/(1 - z^2)$ with center at $z = 1$. We have $1 - z^2 = -(z - 1)(z + 1)$. Using the geometric series

$$\frac{1}{1 - q} = \sum_{n=0}^{\infty} q^n \quad (|q| < 1),$$

we find

$$(a) \quad \begin{aligned} \frac{1}{z+1} &= \frac{1}{2+(z-1)} = \frac{1}{2} \frac{1}{1 - \left(-\frac{z-1}{2}\right)} \\ &= \frac{1}{2} \sum_{n=0}^{\infty} \left(-\frac{z-1}{2}\right)^n = \sum_{n=0}^{\infty} \frac{(-1)^n}{2^{n+1}} (z-1)^n; \end{aligned}$$

this series converges in the disk $|z-1|/2 < 1$, that is, $|z-1| < 2$. Similarly,

$$(b) \quad \begin{aligned} \frac{1}{z+1} &= \frac{1}{(z-1)+2} = \frac{1}{(z-1)} \frac{1}{1 + \frac{2}{z-1}} \\ &= \frac{1}{z-1} \sum_{n=0}^{\infty} \left(-\frac{2}{z-1}\right)^n = \sum_{n=0}^{\infty} \frac{(-2)^n}{(z-1)^{n+1}}; \end{aligned}$$

this series converges for $|2/(z-1)| < 1$, that is, $|z-1| > 2$. Therefore, from (a) we obtain

$$\begin{aligned} f(z) &= \frac{-1}{(z-1)(z+1)} = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{2^{n+1}} (z-1)^{n-1} \\ &= \frac{-1/2}{z-1} + \frac{1}{4} - \frac{1}{8}(z-1) + \frac{1}{16}(z-1)^2 - + \dots; \end{aligned}$$

this series converges in the domain $0 < |z-1| < 2$. From (b),

$$f(z) = - \sum_{n=0}^{\infty} \frac{(-2)^n}{(z-1)^{n+2}} = - \frac{1}{(z-1)^2} + \frac{2}{(z-1)^3} - \frac{4}{(z-1)^4} + - \dots \quad (|z-1| > 2).$$

Example 3. From Prob. 22 at the end of Sec. 14.2 we obtain the Laurent series

$$\cot z = \frac{1}{z} - \frac{1}{3}z - \frac{1}{45}z^3 - \frac{2}{945}z^5 - \dots \quad (0 < |z| < \pi).$$

Problems

Expand the following functions in Laurent series which converge for $0 < |z| < R$ and determine the precise region of convergence.

$$1. \frac{e^{-z}}{z^3} \quad 2. \frac{e^{1/z^2}}{z^6} \quad 3. \frac{\cos 2z}{z^2} \quad 4. \frac{1}{z^4(1+z)}$$

$$5. \frac{1}{z^2(1-z^2)} \quad 6. \frac{1}{z^2(z-3)} \quad 7. \frac{\sinh 3z}{z^3} \quad 8. \frac{1}{z^6+z^4}$$

9. Prove that the Laurent expansion of a given analytic function in a given annulus is unique

Find all Taylor series and Laurent series with center $z = a$ and determine the precise regions of convergence.

10. $\frac{1}{1+z^2}, \quad a = -i$

11. $\frac{1}{z^4}, \quad a = 1$

12. $\frac{1}{z^3}, \quad a = i$

13. $\frac{7z^8 + 9z - 18}{z^3 - 9z}, \quad a = 0$

14. $\frac{4z^3 + 2z - 4}{z^3 - 4z}, \quad a = 2$

15. $\frac{\sin z}{(z - \pi/4)^3}, \quad a = \pi/4$

14.6 Behavior of Functions at Infinity

We shall see in the next section that the Laurent series can be used for classifying the singularities of analytic functions. In this connection we shall also consider the behavior of analytic functions $f(z)$ as $|z|$ approaches infinity.

For investigating a function $f(z)$ for large $|z|$, it seems to be natural to introduce a new variable w by setting $z = 1/w$ because large $|z|$ then corresponds to small $|w|$, and $|z| \rightarrow \infty$ corresponds to $|w| \rightarrow 0$.

The transformation $w = 1/z$ was considered in Sec. 11.4 in the extended complex plane which was obtained by attaching an improper point ("point at infinity") to the finite complex plane. We may then let $z = \infty$ correspond to $w = 0$ (and $w = \infty$ correspond to $z = 0$). The procedure of extending the complex plane may also be motivated by the following simple geometrical consideration.

The usual representation of complex numbers z in the complex plane is convenient as long as the absolute values of the numbers are not too large. For large $|z|$ the situation becomes inconvenient, and in this case we may prefer a representation of the complex numbers on a sphere, which was suggested by Riemann and is obtained as follows.

Let S be a sphere of diameter 1 which touches the complex z -plane at the origin (Fig. 345). Let N be the "North pole" of S (the point diametrically

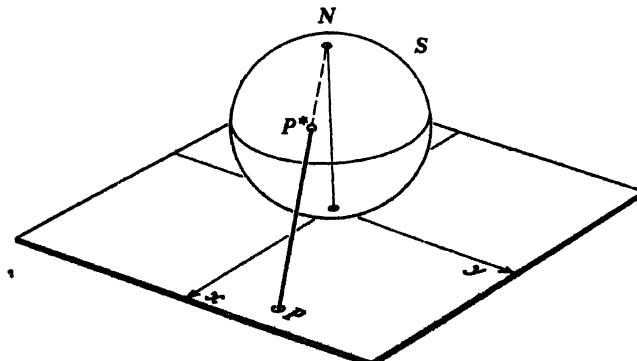


Fig. 345. Riemann number sphere.

opposite to the point of contact between the sphere and the plane). Let P be any point in the complex plane. Then the straight segment with end points P and N intersects S at a point P^* . We let P and P^* correspond to each other. In this way we obtain a correspondence between the points in the complex plane and the points on S , and P^* is the image point of P with respect to this mapping. The complex numbers, first represented in the plane, are now represented by points on S . To each z there corresponds a point on S . Conversely, each point on S represents a complex number z , except for the point N which does not correspond to any point in the complex plane. But if we introduce the improper point $z = \infty$ and let this point correspond to N , the mapping becomes a one-to-one mapping of the extended plane onto S . The sphere S is called the **Riemann number sphere**. The particular mapping which we have used is called a **stereographic projection**.

Obviously, the unit circle is mapped onto the "equator" of S . The interior of the unit circle corresponds to the "Southern hemisphere" and the exterior to the "Northern hemisphere." Numbers z whose absolute values are large lie close to the North pole N . The x and y axes (and, more generally, all the straight lines through the origin) are mapped onto "meridians," while circles with center at the origin are mapped onto "parallels." It can be shown that any circle or straight line in the z -plane is mapped onto a circle on S , and, furthermore, that stereographic projection is conformal.

Given a function $f(z)$ to be investigated for large $|z|$, we may now set $z = 1/w$ and investigate $f(z) = f(1/w) \equiv g(w)$ in a neighborhood of $w = 0$. We define

$$g(0) = \lim_{w \rightarrow 0} g(w),$$

and we say that $f(z)$ is *analytic or singular at infinity according as $g(w)$ is analytic or singular, respectively, at $w = 0$* .

Example 1. The function $f(z) = 1/z^2$ is analytic at infinity since $g(w) = f(1/w) = w^2$ is analytic at $w = 0$. The function $f(z) = z^3$ is singular at infinity because $g(w) = f(1/w) = 1/w^3$ is singular at $w = 0$. The exponential function e^z is singular at infinity since $e^{1/w}$ is singular at $w = 0$. Similarly, the trigonometric functions $\sin z$ and $\cos z$ are singular at infinity.

Let $f(z)$ be analytic outside a circle C , say, in the domain $|z - a| > R$, also at infinity. If we set

$$z = \frac{1}{w} + a, \quad \text{then} \quad z - a = \frac{1}{w},$$

and

$$f(z) = f\left(\frac{1}{w} + a\right) \equiv h(w)$$

is analytic in the disk $|w| < 1/R$. Let the Maclaurin series of $h(w)$ be

$$h(w) = \sum_{n=0}^{\infty} c_n w^n = c_0 + c_1 w + c_2 w^2 + \cdots \quad \left(|w| < \frac{1}{R}\right).$$

Then by inserting $w = 1/(z - a)$ we obtain the Laurent series

$$(1) \quad f(z) = \sum_{n=0}^{\infty} \frac{c_n}{(z-a)^n} = c_0 + \frac{c_1}{z-a} + \frac{c_2}{(z-a)^2} + \cdots \quad (|z-a| > R).$$

Problems

Are the following functions analytic at infinity or not?

- | | | | |
|---------------------|-----------------------|---------------------------|-------------------------|
| 1. $z^2 + z^{-2}$ | 2. $z^{-3} + z^{-1}$ | 3. e^z, e^{z^2}, e^{-z} | 4. $e^{1/z}, e^{1/z^2}$ |
| 5. $\cos z, \sin z$ | 6. $\cosh z, \sinh z$ | 7. $ze^{1/z}$ | 8. $1/(z^2 - 3z)$ |

Describe and sketch the images of the following regions on the Riemann number sphere.

- | | | |
|--|---|---------------------------------|
| 9. $\operatorname{Re} z \leq 0$ | 10. $\operatorname{Im} z \geq 0$ | 11. $ z \geq 5$ |
| 12. $ z \leq 2$ | 13. $1/2 \leq z \leq 2$ | 14. $ z < 3, \arg z < \pi/4$ |
| 15. $-\pi \leq \operatorname{Im} z \leq \pi$ | 16. $ \arg z \leq 3\pi/4$ | 17. $ z > 1, \arg z < \pi/2$ |
| 18. $-\pi/4 < \arg z < 3\pi/4$ | 19. $1 \leq z \leq 5, -\pi/4 \leq \arg z \leq 3\pi/2$ | |
20. Using the method described at the end of this section, show that

$$\frac{1}{z^4} = \sum_{n=0}^{\infty} \binom{-4}{n} (z-1)^{-n-4} = \frac{1}{(z-1)^4} - \frac{4}{(z-1)^5} + \cdots \quad (|z-1| > 1).$$

Chapter 15

Integration by the Method of Residues

Since there are various methods for determining the coefficients of a Laurent series (1), Sec. 14.5, without using the integral formulas (2), Sec. 14.5, we may use the formula for c_1 for evaluating complex integrals in a very elegant and simple fashion. c_1 will be called the *residue* of $f(z)$ at $z = a$. This powerful method may also be applied for evaluating certain real integrals, as we shall see in Sec. 15.4.

Prerequisites for this chapter: Chaps. 10, 12, 14.

References: Appendix 1, Part G.

Answers to problems: Appendix 2.

15.1 Zeros and Singularities

If a function $f(z)$ is analytic in a domain D and is zero at a point $z = a$ in D , then $f(z)$ is said to have a *zero* at that point $z = a$. If not only f but also the derivatives f' , \dots , $f^{(n-1)}$ are all zero at $z = a$ and $f^{(n)}(a) \neq 0$, then $f(z)$ is said to have a *zero of the order n* at the point $z = a$.

For example, if $f(a) = 0$, $f'(a) \neq 0$, then f has a zero of the first order or *simple zero* at $z = a$. If $f(a) = f'(a) = 0$, $f''(a) \neq 0$, then the zero of f at $z = a$ is of the second order, etc.

Furthermore, our consideration in the last section suggests the following definition.

An analytic function $f(z)$ is said to have a *zero of n -th order at infinity* if $f(1/z)$ has such a zero at $z = 0$.

Example 1. The function $\sin z$ has simple zeros at $z = 0, \pm\pi, \pm 2\pi, \dots$. The function $(z - a)^3$ has a zero of the third order at $z = a$. The function $1 - \cos z$ has second-order zeros at $z = 0, \pm 2\pi, \pm 4\pi, \dots$. The function $1/(1 - z)$ has a simple zero at infinity.

Clearly, if $f(z)$ is analytic in some neighborhood of a point $z = a$ and has a zero of n th order at a , it follows from Taylor's theorem in Sec. 14.1 that

the coefficients b_0, \dots, b_{n-1} of its Taylor series with center $z = a$ are zero, and the series is of the form

$$(1) \quad f(z) = b_n(z - a)^n + b_{n+1}(z - a)^{n+1} + \dots \\ = (z - a)^n [b_n + b_{n+1}(z - a) + b_{n+2}(z - a)^2 + \dots] \quad (b_n \neq 0).$$

A point of a point set S is called an *isolated point* of S , if it has a neighborhood which does not contain further points of S . A point b is called a *limit point* of S , if every neighborhood of b (no matter how small) contains at least one point ($\neq b$) of S (and hence infinitely many points of S).

Example 2. The set of the points $z = n$ ($n = 1, 2, \dots$) consists wholly of isolated points and has no limit point in the finite plane. The set of the points $z = i/n$ ($n = 1, 2, \dots$) on the imaginary axis consists wholly of isolated points and has one limit point, namely, $z = 0$; this point does not belong to the set. The set of all complex numbers z for which $|z| < 1$ has no isolated points. Every point of the set and also the points on the unit circle $|z| = 1$ (which do not belong to the set) are limit points of the set.

Theorem 1. *The zeros of an analytic function $f(z)$ ($\not\equiv 0$) are isolated.*

Proof. We consider (1). Let $g(z)$ be the analytic function represented by the series in brackets $[\dots]$. Since $b_n \neq 0$, we have $g(a) \neq 0$. Consequently, since $g(z)$ is continuous, it is not zero in some neighborhood of $z = a$. It follows that $f(z)$ is not zero in that neighborhood (except at $z = a$), so that $z = a$ is the only zero of $f(z)$ in that neighborhood and is, therefore, isolated. This completes the proof.

Analytic functions may have different types of singularities.¹ We first remember that a *singular point* of an analytic function $f(z)$ is a point where $f(z)$ ceases to be analytic (cf. Sec. 14.1). We also say that $f(z)$ is *singular*, or *has a singularity*, at that point. The function $f(z)$ is said to be *singular at infinity*, if $f(1/z)$ is singular at $z = 0$.

If $f(z)$ has an isolated singularity at a point $z = a$, then we can represent it by its **Laurent series** (cf. Sec. 14.5)

$$(2) \quad f(z) = \sum_{n=0}^{\infty} b_n(z - a)^n + \sum_{n=1}^{\infty} \frac{c_n}{(z - a)^n}$$

valid throughout some neighborhood of $z = a$ (except at $z = a$ itself). The last series in (2) is called the **principal part** of $f(z)$ near $z = a$.

It may happen that from some n on, all the coefficients c_n are zero, say, $c_m \neq 0$ and $c_n = 0$ for all $n > m$. Then (2) reduces to the form

$$(3) \quad f(z) = \sum_{n=0}^{\infty} b_n(z - a)^n + \frac{c_1}{z - a} + \dots + \frac{c_m}{(z - a)^m} \quad (c_m \neq 0).$$

In this case, where the principal part consists of finitely many terms, the singularity of f at $z = a$ is called a *pole*, and m is called the *order* of the pole. Poles of the first order are also known as *simple poles*.

¹ By our general assumption the functions under consideration are single-valued.

Any singularity of a single-valued analytic function other than a pole is called an **essential singularity**.

Poles are, by definition, isolated singularities. All singularities which are not isolated (for instance, the limit point of a sequence of poles) are thus essential singularities. An essential singularity may be isolated or not. If in (2) infinitely many c_n are different from zero, then the singularity of $f(z)$ at $z = a$ is not a pole but an isolated essential singularity.

Example 3. The function

$$f(z) = \frac{1}{z(z-2)^5} + \frac{3}{(z-2)^2}$$

has a simple pole at $z = 0$ and a pole of the fifth order at $z = 2$. The functions

$$(4) \quad e^{1/z} = \sum_{n=0}^{\infty} \frac{1}{n! z^n} = 1 + \frac{1}{z} + \frac{1}{2! z^2} + \dots$$

and

$$(5) \quad \sin \frac{1}{z} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)! z^{2n+1}} = \frac{1}{z} - \frac{1}{3! z^3} + \frac{1}{5! z^5} - \dots$$

have an isolated essential singularity at $z = 0$.

The function $\tan(1/z)$ has poles at

$$\frac{1}{z} = \pm \frac{\pi}{2}, \pm \frac{3\pi}{2}, \dots \quad \text{that is,} \quad z = \pm \frac{2}{\pi}, \pm \frac{2}{3\pi}, \dots$$

The limit point $z = 0$ of these points is thus a nonisolated essential singularity of $\tan(1/z)$.

Example 4. The polynomial $f(z) = 2z + 6z^3$ has a pole of the third order at infinity, because $f\left(\frac{1}{z}\right) = \frac{2}{z} + \frac{6}{z^3}$ has such a pole at $z = 0$. More generally, a polynomial of n th degree has a pole of n th order at infinity.

The functions e^z , $\sin z$, $\cos z$ have an isolated essential singularity at infinity, since $e^{1/z}$, $\sin(1/z)$, and $\cos(1/z)$ have an isolated essential singularity at $z = 0$.

A function $f(z)$ which is not analytic at a point $z = a$ but can be made analytic there by assigning some value to $f(z)$ at $z = a$, is said to have a *removable singularity* at $z = a$. Such singularities are not of interest, because they can be removed.

A function which is analytic everywhere in the finite plane is called an **entire function**.

If such a function is also analytic at infinity, it is bounded for all z , and from Liouville's theorem (cf. Sec. 12.6) it follows that it must be a constant. Hence, any entire function which is not a constant must be singular at infinity. For instance, polynomials (of at least the first degree), e^z , $\sin z$, and $\cos z$ are entire functions, and they are singular at infinity.

An analytic function whose only singularities in the finite plane are poles is called a **meromorphic function**.

Example 5. Rational functions with nonconstant denominator, $\tan z$, $\cot z$, $\sec z$, and $\csc z$ are meromorphic functions.

The classification of singularities into poles and essential singularities is not merely a formal matter; the behavior of an analytic function in a neighborhood of an essential singularity is entirely different from that in the neighborhood of a pole.

Example 6. The function $f(z) = 1/z^2$ has a pole at $z = 0$, and $|f| \rightarrow \infty$ as $z \rightarrow 0$ in any manner.

This example illustrates

Theorem 2. If $f(z)$ is analytic and has a pole at $z = a$, then $|f(z)| \rightarrow \infty$ as $z \rightarrow a$ in any manner. (Cf. Prob. 29.)

Example 7. The function $f(z) = e^{1/z}$ has an essential singularity at $z = 0$. It has no limit for approach along the imaginary axis; it becomes infinite if $z \rightarrow 0$ through positive real values, but it approaches zero if $z \rightarrow 0$ through negative real values. It takes on any given value $c = c_0 e^{i\alpha} \neq 0$ in an arbitrarily small neighborhood of $z = 0$. In fact, setting $z = re^{i\theta}$, we have to solve the equation

$$e^{1/z} = e^{(\cos \theta - i \sin \theta)/r} = c_0 e^{i\alpha}$$

for r and θ . Equating the absolute values and the arguments, we have

$$e^{(\cos \theta)/r} = c_0 \quad \text{or} \quad \cos \theta = r \ln c_0 \quad \text{and} \quad \sin \theta = -\alpha r.$$

From these two equations and $\cos^2 \theta + \sin^2 \theta = 1$ we obtain

$$r^2 = \frac{1}{(\ln c_0)^2 + \alpha^2} \quad \text{and} \quad \tan \theta = -\frac{\alpha}{\ln c_0}$$

We see that r can be made arbitrarily small by adding multiples of 2π to α , leaving c unaltered.

This example illustrates the famous

Picard's theorem. If $f(z)$ is analytic and has an isolated essential singularity at a point a , it takes on every value, with at most one exceptional value, in an arbitrarily small neighborhood of a .

In Ex. 7, the exceptional value is $z = 0$. The proof of Picard's theorem is rather complicated; it can be found in Ref. [G9].

Problems

Determine the location and order of the zeros of the following functions.

- | | | | |
|-------------------|---------------------|-----------------------|-----------------------|
| 1. $z^3 - 1$ | 2. z^3 | 3. $(z + i)^3$ | 4. $\cos^2 z$ |
| 5. $\sin^3 \pi z$ | 6. $(\sin z - 1)^3$ | 7. $\cosh^2 z$ | 8. $z^2 e^z$ |
| 9. $e^z - e^{iz}$ | 10. $\cot^2 z$ | 11. $z^4 \sin^2(1/z)$ | 12. $z^{-5}(z - 6)^2$ |

Determine the location and type of singularities of the following functions.

- | | | |
|------------------------------|--------------------|-----------------------|
| 13. $z + z^{-1}$ | 14. $z^3 + z^{-2}$ | 15. $z^3 + az + b$ |
| 16. $1/(1 - z)$ | 17. $1/(z + a)^3$ | 18. $\cosh z$ |
| 19. $e^z + e^{1/z}$ | 20. $\tan \pi z$ | 21. $\cos z - \sin z$ |
| 22. $(\cos z - \sin z)^{-1}$ | 23. $\sin(1/z)$ | 24. e^z/z^5 |
| 25. $e^{1/z}/(z - 1)$ | 26. $(e^z - 1)/z$ | 27. $e^z/(z^3 + 1)^2$ |

28. If $f(z)$ has a zero of order n at $z = a$, prove that $f''(z)$ has a zero of order $2n$ at $z = a$. Give examples.
29. Verify Theorem 2 for $f(z) = z^{-2} + z^{-1}$. Prove Theorem 2.
30. If $f(z)$ has a zero of order n at $z = a$, prove that $1/f(z)$ has a pole of order n at $z = a$. Give examples.

15.2 Residues

If $f(z)$ is analytic in a neighborhood of a point $z = a$, then, by Cauchy's integral theorem,

$$(1) \quad \int_C f(z) dz = 0$$

for any contour in that neighborhood. If, however, $f(z)$ has a pole or an isolated essential singularity (cf. Sec. 15.1) at $z = a$ and a lies in the interior of C , then the integral in (1) will, in general, be different from zero. In this case we may represent $f(z)$ by a Laurent series

$$(2) \quad f(z) = \sum_{n=0}^{\infty} b_n(z - a)^n + \sum_{n=1}^{\infty} \frac{c_n}{(z - a)^n}$$

which converges in the domain $0 < |z - a| < R$, where R is the distance from a to the nearest singular point of $f(z)$. From (2) in Sec. 14.5 we see that the coefficient c_1 of the power $1/(z - a)$ in this development is

$$c_1 = \frac{1}{2\pi i} \int_C f(z) dz,$$

and, therefore,

$$(3) \quad \int_C f(z) dz = 2\pi i c_1.$$

the integration being taken in the counterclockwise sense around a simple closed path C which lies in the domain $0 < |z - a| < R$ and contains the point $z = a$ in its interior. The coefficient c_1 in the development (2) of $f(z)$ is called the residue of $f(z)$ at $z = a$, and we shall use the notation

$$(4) \quad c_1 = \underset{z=a}{\text{Res}} f(z).$$

We have seen that Laurent expansions can be obtained by various methods, without using the integral formulas for the coefficients. Hence, we may determine the residue by one of those methods and then use the formula (3) for evaluating contour integrals.

Example 1. Integrate the function $f(z) = z^{-4} \sin z$ around the unit circle C in the counterclockwise sense.

From (5) in Sec. 14.2 we obtain the Laurent series

$$f(z) = \frac{\sin z}{z^4} = \frac{1}{z^3} - \frac{1}{3!z} + \frac{z}{5!} - \frac{z^3}{7!} + \cdots$$

We see that the residue is $c_1 = -1/3!$, and from (3) it follows that

$$\int_C \frac{\sin z}{z^4} dz = 2\pi i c_1 = -\frac{\pi i}{3}.$$

Before we proceed in evaluating integrals we shall develop a simple standard method for determining the residue in the case of a pole.

If $f(z)$ has a simple pole at a point $z = a$, the corresponding Laurent series is of the form [cf. (3) in the last section]

$$f(z) = \frac{c_1}{z-a} + b_0 + b_1(z-a) + b_2(z-a)^2 + \dots \quad (0 < |z-a| < R, c_1 \neq 0).$$

Multiplying both sides by $z-a$, we have

$$(5) \quad (z-a)f(z) = c_1 + (z-a)[b_0 + b_1(z-a) + \dots].$$

If we let z approach a , the right-hand side approaches c_1 and we obtain

$$(6) \quad \text{Res } f(z) = c_1 = \lim_{z \rightarrow a} (z-a)f(z).$$

In the case of a simple pole, another useful formula is obtained as follows. If $f(z)$ has a simple pole at $z = a$, we may set

$$f(z) = \frac{p(z)}{q(z)}$$

where $p(z)$ and $q(z)$ are analytic at $z = a$, $p(a) \neq 0$, and $q(z)$ has a simple zero at $z = a$. Consequently, $q(z)$ can be expanded in a Taylor series of the form

$$q(z) = (z-a)q'(a) + \frac{(z-a)^2}{2!} q''(a) + \dots$$

From (6) it follows that

$$\text{Res } f(z) = \lim_{z \rightarrow a} (z-a) \frac{p(z)}{q(z)} = \lim_{z \rightarrow a} \frac{(z-a)p(z)}{(z-a)[q'(a) + (z-a)q''(a)/2 + \dots]} ;$$

that is, in the case of a simple pole we also have

$$(7) \quad \text{Res } f(z) = \text{Res}_{z=a} \frac{p(z)}{q(z)} = \frac{p(a)}{q'(a)} .$$

Example 2. The function $f(z) = (4-3z)/(z^2-z)$ has simple poles at $z = 0$ and $z = 1$. From (7) we obtain

$$\text{Res } f(z) = \left[\frac{4-3z}{2z-1} \right]_{z=0} = -4, \quad \text{Res } f(z) = \left[\frac{4-3z}{2z-1} \right]_{z=1} = 1.$$

We shall now consider poles of higher order. If $f(z)$ has a pole of order $m > 1$ at a point $z = a$, the corresponding Laurent expansion is of the form

$$f(z) = \frac{c_m}{(z-a)^m} + \frac{c_{m-1}}{(z-a)^{m-1}} + \dots + \frac{c_2}{(z-a)^2} + \frac{c_1}{z-a} + b_0 + b_1(z-a) + \dots$$

where $c_m \neq 0$ and the series converges in some neighborhood of $z = a$, except at the point a itself. By multiplying both sides by $(z - a)^m$ we obtain

$$(z - a)^m f(z) = c_m + c_{m-1}(z - a) + \cdots + c_2(z - a)^{m-2} + c_1(z - a)^{m-1} \\ + b_0(z - a)^m + b_1(z - a)^{m+1} + \cdots.$$

This shows that the residue c_1 of $f(z)$ at $z = a$ is now the coefficient of the power $(z - a)^{m-1}$ in the Taylor expansion of the function $g(z) = (z - a)^m f(z)$ with center at $z = a$. Therefore, from Taylor's theorem (Sec. 14.1),

$$c_1 = \frac{g^{(m-1)}(a)}{(m-1)!};$$

that is,

$$(8) \quad \text{Res } f(z) = \frac{1}{(m-1)!} \lim_{z \rightarrow a} \left\{ \frac{d^{m-1}}{dz^{m-1}} [(z-a)^m f(z)] \right\}.$$

Example 3. The function

$$f(z) = \frac{2z}{(z+4)(z-1)^2}$$

has a pole of the second order at $z = 1$, and from (8) we obtain the corresponding residue

$$\text{Res } f(z) = \lim_{z \rightarrow 1} \frac{d}{dz} [(z-1)^2 f(z)] = \lim_{z \rightarrow 1} \frac{d}{dz} \left(\frac{2z}{z+4} \right) = \frac{8}{25}.$$

Of course, in the case of a rational function $f(z)$ the residues can also be determined from the representation of $f(z)$ in terms of partial fractions.

Problems

Find the residues at the singular points of the following functions.

1. $\frac{1}{1-z}$

2. $\frac{z+3}{z+1}$

3. $\frac{1}{z^2}$

4. $\frac{1}{1+z^2}$

5. $\frac{1}{z^2-1}$

6. $\frac{1}{(z^2-1)^2}$

7. $\frac{z}{z^4-1}$

8. $\frac{1}{(z^4-1)^2}$

9. $\csc z$

10. $\sec z$

11. $\cot z$

12. $\tan z$

13. e^z/z

14. e^{z^2}/z^2

15. $e^{1/z}/z$

16. $e^{1/z}/z^2$

17. $z^{-2} \sin z$

18. $z^{-4} \sin^2 z$

19. $\csc^2 z$

20. $1/(1-e^z)$

21. $\frac{z+1}{(z^2-16)(z+2)}$

22. $\frac{4-3z}{z^3-3z^2+2z}$

23. $\frac{\cos z}{z^2-4\pi z+3\pi^2}$

24. $\frac{4z-1}{z^2+3z+2}$

25. $\frac{\cos z}{z^6}$

26. $\frac{e^z}{(z-\pi i)^5}$

Evaluate the following integrals where C is the unit circle (counterclockwise).

27. $\int_C e^{1/z} dz$

28. $\int_C z e^{1/z} dz$

29. $\int_C z^2 e^{1/z} dz$

30. $\int_C z^{-2} e^{1/z} dz$

31. $\int_C \cot z dz$

32. $\int_C \tan z dz$

33. $\int_C \frac{dz}{\sin z}$

34. $\int_C \frac{dz}{\sinh z}$

35. $\int_C \frac{dz}{\cosh z}$

36. $\int_C \frac{z}{2z+i} dz$

37. $\int_C \frac{z^2 - 4}{(z-2)^4} dz$

38. $\int_C \frac{z^3 - 3z}{4z^2 + 8z + 3} dz$

15.3 The Residue Theorem

So far we are in a position to evaluate contour integrals whose integrands have only one isolated singularity inside the contour of integration. We shall now see that our simple method may easily be extended to the case when the integrand has several isolated singularities inside the contour.

Residue theorem. *Let $f(z)$ be a single-valued function which is analytic inside a simple closed path C and on C , except for finitely many singular points a_1, a_2, \dots, a_m inside C . Then*

(1)
$$\int_C f(z) dz = 2\pi i \sum_{j=1}^m \operatorname{Res}_{z=a_j} f(z),$$

the integral being taken in the counterclockwise sense around C .

Proof. We enclose each of the singular points a_j in a circle C_j , with radius small enough that those m circles and C are all separated (Fig. 346). Then $f(z)$ is analytic in the multiply connected domain D bounded by C and C_1, \dots, C_m and on the entire boundary of D . From Cauchy's integral theorem it follows that

(2)
$$\int_C f(z) dz + \int_{C_1} f(z) dz + \int_{C_2} f(z) dz + \cdots + \int_{C_m} f(z) dz = 0,$$

the integral along C being taken in the counterclockwise sense and the other integrals in the clockwise sense (cf. p. 600). We now reverse the sense of integration along C_1, \dots, C_m . Then the signs of the values of these integrals change, and we obtain from (2)

(3)
$$\int_C f(z) dz = \int_{C_1} f(z) dz + \int_{C_2} f(z) dz + \cdots + \int_{C_m} f(z) dz,$$

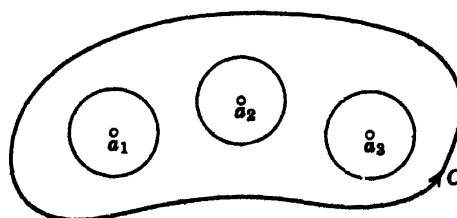


Fig. 346. Residue theorem.

all integrals now being taken in the counterclockwise sense. Since, by (3) in the last section,

$$\int_C f(z) dz = 2\pi i \operatorname{Res}_{z=a} f(z),$$

formula (3) yields (1), and the theorem is proved.

This important theorem has various applications in connection with complex and real integrals. We shall first consider some complex integrals.

Example 1. The function $(4 - 3z)/(z^2 - z)$ is analytic except at the points 0 and 1 where it has simple poles; the residues are -4 and 1 , respectively (cf. Ex. 2 in the last section). Therefore,

$$\int_C \frac{4 - 3z}{z^2 - z} dz = 2\pi i(-4 + 1) = -6\pi i$$

for every simple closed path C which encloses the points 0 and 1, and

$$\int_C \frac{4 - 3z}{z^2 - z} dz = 2\pi i(-4) = -8\pi i$$

for any simple closed path C for which $z = 0$ lies inside C and $z = 1$ lies outside C , the integrations being taken in the counterclockwise sense.

Example 2. Integrate $1/(z^3 - 1)^2$ in the counterclockwise sense around the circle $|z - 1| = 1$. The function has poles of second order at $z = 1$, $z = e^{2\pi i/3}$, and $z = e^{-2\pi i/3}$. Only the pole at $z = 1$ lies inside C . Using (8), Sec. 15.2, we thus obtain

$$\int_C \frac{dz}{(z^3 - 1)^2} = 2\pi i \operatorname{Res}_{z=1} \frac{1}{(z^3 - 1)^2} = 2\pi i \left(-\frac{2}{9} \right) = -\frac{4\pi i}{9}.$$

Example 3. Integrate $1/(z - a)^m$ (m a positive integer) in the counterclockwise sense around any simple closed path C enclosing the point $z = a$. We find

$$\operatorname{Res}_{z=a} \frac{1}{z - a} = 1 \quad \text{and} \quad \operatorname{Res}_{z=a} \frac{1}{(z - a)^m} = 0 \quad (m = 2, 3, \dots).$$

Therefore, the result is

$$\int_C \frac{dz}{(z - a)^m} = \begin{cases} 2\pi i & (m = 1) \\ 0 & (m = 2, 3, \dots), \end{cases}$$

in agreement with Ex. 3 in Sec. 12.1.

Problems

Integrate $\frac{3z^2 + 2z - 4}{z^3 - 4z}$ around the following paths C in the counterclockwise sense.

1. $|z| = 1$
2. $|z| = 3$
3. $|z - 4| = 1$
4. $|z - 4| = 3$
5. $|z - 4| = 5$
6. The boundary of the triangle with vertices $1 - i$, $-1 + 2i$, $-3 - i$

Evaluate the following integrals where C is the unit circle (counterclockwise).

7. $\int_C \frac{3z}{3z - 1} dz$

8. $\int_C \frac{z}{4z^2 - 1} dz$

9. $\int_C \frac{dz}{z^2 - 2z}$

10. $\int_C \frac{dz}{z^2 + 4}$

11. $\int_C \frac{z+1}{4z^3 - z} dz$

12. $\int_C \frac{z^5 - 3z^3 + 1}{(2z+1)(z^2 + 4)} dz$

13. $\int_C \frac{z}{1 + 9z^2} dz$

14. $\int_C \frac{z+1}{z^4 - 2z^3} dz$

15. $\int_C \frac{(z+4)^3}{z^4 + 5z^3 + 6z^2} dz$

16. $\int_C \tan z dz$

17. $\int_C \tan \pi z dz$

18. $\int_C \frac{6z^3 - 4z + 1}{(z-2)(1+4z^2)} dz$

19. $\int_C \tan 2\pi z dz$

20. $\int_C \frac{\tan \pi z}{z^3} dz$

21. $\int_C \frac{e^z}{z^2 - 5z} dz$

22. $\int_C \frac{e^z}{\sin z} dz$

23. $\int_C \frac{e^z}{\cos z} dz$

24. $\int_C \frac{e^z}{\cos \pi z} dz$

25. $\int_C \frac{\cosh z}{z^2 - 3iz} dz$

26. $\int_C \coth z dz$

27. $\int_C \frac{\sinh z}{2z-i} dz$

28. $\int_C \cot z dz$

29. $\int_C \frac{\cot z}{z} dz$

30. $\int_C \frac{e^{z^2}}{\cos \pi z} dz$

5.4 Evaluation of Real Integrals

The residue theorem yields a very elegant and simple method for evaluating certain classes of complicated real integrals, and we shall see that in many cases this method of integration can be made a routine matter.

A. Integrals of rational functions of $\cos \theta$ and $\sin \theta$. We first consider integrals of the type

(1)
$$I = \int_0^{2\pi} R(\cos \theta, \sin \theta) d\theta$$

where $R(\cos \theta, \sin \theta)$ is a real rational function of $\cos \theta$ and $\sin \theta$ finite on the interval $0 \leq \theta \leq 2\pi$. Setting $e^{i\theta} = z$, we obtain

$$\cos \theta = \frac{1}{2}(e^{i\theta} + e^{-i\theta}) = \frac{1}{2}\left(z + \frac{1}{z}\right),$$

$$\sin \theta = \frac{1}{2i}(e^{i\theta} - e^{-i\theta}) = \frac{1}{2i}\left(z - \frac{1}{z}\right),$$

and we see that the integrand becomes a rational function of z , say, $f(z)$. As θ ranges from 0 to 2π , the variable z ranges around the unit circle $|z| = 1$ in the counterclockwise sense. Since $dz/d\theta = ie^{i\theta}$, we have $d\theta = dz/iz$, and the given integral takes the form

(2)
$$I = \int_C f(z) \frac{dz}{iz},$$

the integration being taken in the counterclockwise sense around the unit circle.

Example 1. Let p be a fixed real number in the interval $0 < p < 1$. We consider

$$\int_0^{2\pi} \frac{d\theta}{1 - 2p \cos \theta + p^2} = \int_C \frac{dz/iz}{1 - 2p \frac{1}{2}\left(z + \frac{1}{z}\right) + p^2} = \int_C \frac{dz}{i(1 - pz)(z - p)}.$$

The integrand has simple poles at $z = 1/p > 1$ and $z = p < 1$. Only the last pole lies inside the unit circle C , and the residue is

$$\text{Res}_{z=p} \frac{1}{i(1 - pz)(z - p)} = \left[\frac{1}{i(1 - pz)} \right]_{z=p} = \frac{1}{i(1 - p^2)}.$$

The residue theorem yields

$$\int_0^{2\pi} \frac{d\theta}{1 - 2p \cos \theta + p^2} = 2\pi i \frac{1}{i(1 - p^2)} = \frac{\pi}{1 - p^2} \quad (0 < p < 1).$$

B. Improper integrals of rational functions. We consider real integrals of the type

$$(3) \quad \int_{-\infty}^{\infty} f(x) dx.$$

Such an integral, for which the interval of integration is not finite, is called an **improper integral**, and it has the meaning

$$(4') \quad \int_{-\infty}^{\infty} f(x) dx = \lim_{a \rightarrow -\infty} \int_a^0 f(x) dx + \lim_{b \rightarrow \infty} \int_0^b f(x) dx.$$

If both limits exist, we may couple the two independent passages to $-\infty$ and ∞ , and write²

$$(4) \quad \int_{-\infty}^{\infty} f(x) dx = \lim_{r \rightarrow \infty} \int_{-r}^r f(x) dx.$$

We assume that the function $f(x)$ in (3) is a real rational function whose denominator is different from zero for all real x and is of degree at least two units higher than the degree of the numerator. Then the limits in (4') exist, and we may start from (4). We consider the corresponding contour integral

$$(4*) \quad \int_C f(z) dz$$

around a path C , as shown in Fig. 347 on p. 682. Since $f(x)$ is rational, $f(z)$ has finitely many poles in the upper half-plane, and if we choose r large enough, C encloses all these poles. By the residue theorem we then obtain

$$\int_C f(z) dz = \int_S f(z) dz + \int_{-r}^r f(x) dx = 2\pi i \sum \text{Res } f(z)$$

² The expression on the right side of (4) is called the *Cauchy principal value* of the integral; it may exist even if the limits in (4') do not exist. For instance,

$$\lim_{r \rightarrow \infty} \int_{-r}^r x dx = \lim_{r \rightarrow \infty} \left(\frac{r^2}{2} - \frac{r^2}{2} \right) = 0, \quad \text{but} \quad \lim_{b \rightarrow \infty} \int_0^b x dx = \infty.$$



Fig. 347. Path of the contour integral in (4*).

where the sum consists of all the residues of $f(z)$ at the points in the upper half-plane where $f(z)$ has a pole. From this we have

$$(5) \quad \int_{-r}^r f(x) dx = 2\pi i \sum \text{Res } f(z) - \int_S f(z) dz.$$

We prove that, if $r \rightarrow \infty$, the value of the integral over the semicircle S approaches zero. If we set $z = re^{i\theta}$, then S is represented by $r = \text{const}$, and as z ranges along S the variable θ ranges from 0 to π . Since the degree of the denominator of $f(z)$ is at least two units higher than the degree of the numerator, we have

$$|f(z)| < \frac{k}{|z|^2} \quad (|z| = r > r_0)$$

for sufficiently large constants k and r_0 . By applying (4) in Sec. 12.2 we thus obtain

$$\left| \int_S f(z) dz \right| < \frac{k}{r^2} \pi r = \frac{k\pi}{r} \quad (r > r_0).$$

Hence, as r approaches infinity, the value of the integral over S approaches zero, and (4) and (5) yield the result

$$(6) \quad \int_{-\infty}^{\infty} f(x) dx = 2\pi i \sum \text{Res } f(z),$$

the sum being extended over all the residues of $f(z)$ corresponding to poles of $f(z)$ in the upper half-plane.

Example 2. Using (6), let us show that

$$\int_0^{\infty} \frac{dx}{1+x^4} = \frac{\pi}{2\sqrt{2}}.$$

Indeed, $f(z) = 1/(1+z^4)$ has four simple poles at the points

$$z_1 = e^{\pi i/4}, \quad z_2 = e^{3\pi i/4}, \quad z_3 = e^{-3\pi i/4}, \quad z_4 = e^{-\pi i/4}.$$

The first two of these poles lie in the upper half-plane (Fig. 348). From (7) in Sec. 15.2 we find

$$\text{Res } f(z) = \left[\frac{1}{(1+z^4)} \right]_{z=z_1} = \left[\frac{1}{4z^3} \right]_{z=z_1} = \frac{1}{4} e^{-3\pi i/4} = -\frac{1}{4} e^{\pi i/4},$$

$$\text{Res } f(z) = \left[\frac{1}{4z^3} \right]_{z=z_2} = \frac{1}{4} e^{-\pi i/4} = \frac{1}{4} e^{-\pi i/4}.$$

By (1) in Sec. 10.6 and (6) in the current section,

$$(7) \quad \int_{-\infty}^{\infty} \frac{dx}{1+x^4} = \frac{2\pi i}{4} (-e^{\pi i/4} + e^{-\pi i/4}) = \pi \sin \frac{\pi}{4} = \frac{\pi}{\sqrt{2}}.$$

Since $1/(1+x^4)$ is an even function,

$$\int_0^{\infty} \frac{dx}{1+x^4} = \frac{1}{2} \int_{-\infty}^{\infty} \frac{dx}{1+x^4}.$$

From this and (7) the result follows.

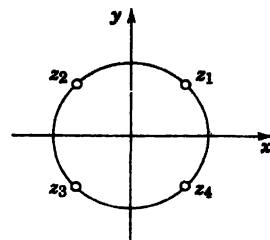


Fig. 348. Example 2.

C. Fourier integrals. Real integrals of the form

$$(8) \quad \int_{-\infty}^{\infty} f(x) \cos sx dx \quad \text{and} \quad \int_{-\infty}^{\infty} f(x) \sin sx dx \quad (s \text{ real})$$

occur in connection with the Fourier integral (cf. Sec. 8.10).

If $f(x)$ is a rational function satisfying the assumptions stated in connection with (3), then the integrals (8) may be evaluated in a similar way as the integrals (3).

In fact, we may then consider the corresponding intégral

$$\int_C f(z) e^{isz} dz \quad (s \text{ real and positive})$$

over the contour C in Fig. 347, and instead of (6) we now obtain the formula

$$(9) \quad \int_{-\infty}^{\infty} f(x) e^{isx} dx = 2\pi i \sum \operatorname{Res} [f(z) e^{isz}] \quad (s > 0)$$

where the sum consists of the residues of $f(z)e^{isz}$ at its poles in the upper half-plane. Equating the real and the imaginary parts on both sides of (9), we have

$$(10) \quad \begin{aligned} \int_{-\infty}^{\infty} f(x) \cos sx dx &= -2\pi \sum \operatorname{Im} \operatorname{Res} [f(z) e^{isz}], \\ \int_{-\infty}^{\infty} f(x) \sin sx dx &= 2\pi \sum \operatorname{Re} \operatorname{Res} [f(z) e^{isz}]. \end{aligned} \quad (s > 0)$$

We remember that (6) was established by proving that the value of the integral over the semicircle S in Fig. 347 approaches zero as $r \rightarrow \infty$. To establish (9) we should now prove the same fact for our present contour integral. This can be done as follows. Since S lies in the upper half-plane $y \geq 0$ and $s > 0$, we obtain

$$|e^{isz}| = |e^{isx}| |e^{-sy}| = e^{-sy} \leq 1 \quad (s > 0, y \geq 0).$$

From this we obtain the inequality

$$|f(z) e^{isz}| = |f(z)| |e^{isz}| \leq |f(z)| \quad (s > 0, y \geq 0)$$

which reduces our present problem to that in part *B* of this section. Continuing as before, we see that the value of the integral under consideration approaches zero as r approaches infinity, and (9) is now established.

Example 3. Using (10), let us show that

$$\int_{-\infty}^{\infty} \frac{\cos sx}{k^2 + x^2} dx = \frac{\pi}{k} e^{-ks}, \quad \int_{-\infty}^{\infty} \frac{\sin sx}{k^2 + x^2} dx = 0 \quad (s > 0, k > 0).$$

In fact, $e^{isz}/(k^2 + z^2)$ has only one pole in the upper half-plane, namely, a simple pole at $z = ik$, and from (7) in Sec. 15.2 we obtain

$$\text{Res}_{z=ik} \frac{e^{isz}}{k^2 + z^2} = \left[\frac{e^{isz}}{2z} \right]_{z=ik} = \frac{e^{-ks}}{2ik}.$$

Therefore,

$$\int_{-\infty}^{\infty} \frac{e^{isz}}{k^2 + x^2} dx = 2\pi i \frac{e^{-ks}}{2ik} = \frac{\pi}{k} e^{-ks}$$

and this yields the above results [cf. also (14) in Sec. 8.10].

¶ **D. Other types of real improper integrals** are definite integrals

$$(11) \quad \int_A^B f(x) dx$$

whose integrand becomes infinite at a point a in the interval of integration, that is,

$$\lim_{x \rightarrow a} |f(x)| = \infty.$$

Then the integral (11) means

$$(12) \quad \int_A^B f(x) dx = \lim_{\epsilon \rightarrow 0} \int_A^{a-\epsilon} f(x) dx + \lim_{\eta \rightarrow 0} \int_{a+\eta}^B f(x) dx$$

where both ϵ and η approach zero independently and through positive values. It may happen that neither of these limits exists when $\epsilon, \eta \rightarrow 0$ independently, but

$$(13) \quad \lim_{\epsilon \rightarrow 0} \left[\int_A^{a-\epsilon} f(x) dx + \int_{a+\epsilon}^B f(x) dx \right]$$

exists; this is called the **Cauchy principal value** of the integral and is often written

$$\text{pr. v. } \int_A^B f(x) dx.$$

For example,

$$\text{pr. v. } \int_{-1}^1 \frac{dx}{x^3} = \lim_{\epsilon \rightarrow 0} \left[\int_{-1}^{-\epsilon} \frac{dx}{x^3} + \int_{\epsilon}^1 \frac{dx}{x^3} \right] = 0;$$

the principal value exists although the integral itself has no meaning. The whole situation is quite similar to that discussed in part *B* of this section.

To evaluate improper integrals whose integrands have poles on the real axis, we use a path which avoids these singularities by following small semicircles with centers at the singular points; the procedure may be illustrated by the following example.

Example 4. Show that

$$\int_0^\infty \frac{\sin x}{x} dx = \frac{\pi}{2} \quad (\text{cf. Sec. 8.10}).$$

We do not consider $(\sin z)/z$ because this function does not behave suitably at infinity. We consider e^{iz}/z , which has a simple pole at $z = 0$, and integrate around the contour shown in Fig. 349. Since e^{iz}/z is analytic inside and on C , it follows from Cauchy's integral theorem that

$$(14) \quad \int_C \frac{e^{iz}}{z} dz = 0.$$

We show that the value of the integral over the large semicircle C_1 approaches zero as R approaches infinity. Setting $z = Re^{i\theta}$ we have $dz = iRe^{i\theta} d\theta$, $dz/z = i d\theta$ and therefore

$$\left| \int_{C_1} \frac{e^{iz}}{z} dz \right| = \left| \int_0^\pi e^{iz} i d\theta \right| \leq \int_0^\pi |e^{iz}| d\theta \quad (z = Re^{i\theta}).$$

In the integrand on the right,

$$|e^{iz}| = |e^{iR(\cos \theta + i \sin \theta)}| = |e^{iR \cos \theta}| |e^{-R \sin \theta}| = e^{-R \sin \theta}.$$

By inserting this and using the fact that in the interval of integration,

$$\sin(\pi - \theta) = \sin \theta$$

we obtain

$$\int_0^\pi |e^{iz}| d\theta = \int_0^\pi e^{-R \sin \theta} d\theta = 2 \int_0^{\pi/2} e^{-R \sin \theta} d\theta = 2 \left[\int_0^\epsilon e^{-R \sin \theta} d\theta + \int_\epsilon^{\pi/2} e^{-R \sin \theta} d\theta \right]$$

where ϵ is any value between 0 and $\pi/2$. The absolute value of the integrand in the first and the last integral on the right is at most equal to 1 and $e^{-R \sin \epsilon}$, respectively, because the integrand is a monotone decreasing function of θ in the interval of integration. Consequently, the whole expression on the right is smaller than

$$2 \left[\int_0^\epsilon d\theta + e^{-R \sin \epsilon} \int_\epsilon^{\pi/2} d\theta \right] = 2 \left[\epsilon + e^{-R \sin \epsilon} \left(\frac{\pi}{2} - \epsilon \right) \right] < 2\epsilon + \pi e^{-R \sin \epsilon}.$$

Altogether,

$$\left| \int_{C_1} \frac{e^{iz}}{z} dz \right| < 2\epsilon + \pi e^{-R \sin \epsilon}.$$

We first take ϵ arbitrarily small. Then, having fixed ϵ , the last term can be made as small as we please by choosing R sufficiently large. Hence, the value of the integral along C_1 approaches zero as R approaches infinity.

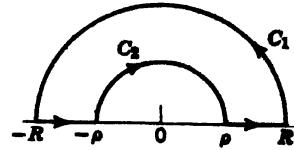


Fig. 349. Example 4.

For the integral over the small semicircle C_2 in Fig. 349 we have

$$\int_{C_2} \frac{e^{iz}}{z} dz = \int_{C_2} \frac{dz}{z} + \int_{C_2} \frac{e^{iz} - 1}{z} dz.$$

The first integral on the right has the value $-\pi i$. The integrand of the last integral is analytic at $z = 0$ and, therefore, bounded in absolute value as $\rho \rightarrow 0$. From this and (4) in Sec. 12.2 we conclude that the value of that integral approaches zero as $\rho \rightarrow 0$. From (14) we thus obtain

$$\text{pr. v. } \int_{-\infty}^{\infty} \frac{e^{ix}}{x} dx = -\lim_{\rho \rightarrow 0} \int_{C_2} \frac{e^{iz} - 1}{z} dz = +\pi i,$$

and by taking the imaginary parts on both sides

$$(15) \quad \text{pr. v. } \int_{-\infty}^{\infty} \frac{\sin x}{x} dx = \pi.$$

Now the integrand in (15) is not singular at $x = 0$. Furthermore, since for positive x the function $1/x$ decreases, the areas under the curve of the integrand between two consecutive positive zeros decrease in a monotone fashion, that is, the absolute values of the integrals

$$I_n = \int_{n\pi}^{(n+1)\pi} \frac{\sin x}{x} dx \quad n = 0, 1, \dots$$

form a monotone decreasing sequence $|I_0|, |I_1|, \dots$, and $I_n \rightarrow 0$ as $n \rightarrow \infty$. Since these integrals have alternating sign, it follows from the Leibniz criterion in Sec. 13.2 that the series $I_0 + I_1 + I_2 + \dots$ converges. Clearly, the sum of the series is the integral

$$\int_0^{\infty} \frac{\sin x}{x} dx = \lim_{b \rightarrow \infty} \int_0^b \frac{\sin x}{x} dx,$$

which therefore exists. Similarly, the integral from 0 to $-\infty$ exists. Hence we need not take the principal value in (15), and

$$\int_{-\infty}^{\infty} \frac{\sin x}{x} dx = \pi.$$

Since the integrand is an even function, the desired result follows.

E. Reduction of integrals to other known integrals. By means of contour integration, certain real improper integrals may be reduced to other real integrals whose values are known. This may be illustrated by

Example 5. We shall see in Sec. 17.2 that

$$(16) \quad \int_0^{\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

Using this formula show that

$$(17) \quad \int_0^{\infty} \cos(x^2) dx = \frac{1}{2} \sqrt{\frac{\pi}{2}}, \quad \int_0^{\infty} \sin(x^2) dx = \frac{1}{2} \sqrt{\frac{\pi}{2}}$$

(These integrals are of importance in connection with the so-called Fresnel integrals to be considered in Sec. 17.2.)

To prove (17), we integrate the function e^{-z^2} along the contour shown in Fig. 350. By Cauchy's integral theorem,

$$(18) \quad \int_C e^{-z^2} dz = \int_0^R e^{-x^2} dx + \int_{C_1} e^{-z^2} dz + \int_S e^{-z^2} dz = 0.$$

We prove that the value of the integral along C_1 approaches zero as R approaches infinity. Setting $z = Re^{i\theta}$, we have

$$\begin{aligned} \int_{C_1} e^{-z^2} dz &= \left| \int_0^{\pi/4} e^{-R^2(\cos 2\theta + i \sin 2\theta)} iRe^{i\theta} d\theta \right| \\ &\leq R \int_0^{\pi/4} e^{-R^2 \cos 2\theta} d\theta < R \int_0^{\pi/4} e^{-R^2(1 - 4\theta/\pi)} d\theta = \frac{\pi}{4R} (1 - e^{-R^2}), \end{aligned}$$

where the last inequality follows from

$$\cos 2\theta \geq 1 - \frac{4\theta}{\pi} \quad \left(0 \leq \theta \leq \frac{\pi}{4} \right).$$

If $R \rightarrow \infty$, the expression on the right approaches zero.

Since the segment S may be represented in the form

$$z(t) = te^{i\pi/4} \quad (0 \leq t \leq R),$$

we have $z^2 = it^2$, $dz = e^{i\pi/4} dt$, and

$$\int_S e^{-z^2} dz = e^{i\pi/4} \int_R^0 e^{-it^2} dt.$$

From (18) and (16) we thus obtain

$$\int_0^\infty e^{-it^2} dt = \lim_{R \rightarrow \infty} \int_0^R e^{-it^2} dt = e^{-i\pi/4} \lim_{R \rightarrow \infty} \int_0^R e^{-x^2} dx = \frac{1}{2}\sqrt{\pi} e^{-i\pi/4} = \frac{1}{2}\sqrt{\frac{\pi}{2}}(1-i).$$

If we separate the real and the imaginary parts, (17) follows.

Problems

Evaluate the following integrals.

$$1. \int_0^{2\pi} \frac{d\theta}{2 + \cos \theta}$$

$$2. \int_0^{\pi} \frac{d\theta}{1 + k \cos \theta} \quad (k^2 < 1)$$

$$3. \int_0^{\pi} \frac{d\theta}{k + \cos \theta} \quad (k > 1)$$

$$4. \int_0^{2\pi} \frac{d\theta}{25 - 24 \cos \theta}$$

$$5. \int_0^{2\pi} \frac{d\theta}{5 - 3 \cos \theta}$$

$$6. \int_0^{2\pi} \frac{\cos \theta}{17 - 8 \cos \theta} d\theta$$

$$7. \int_0^{2\pi} \frac{\cos \theta}{3 + \sin \theta} d\theta$$

$$8. \int_0^{2\pi} \frac{\cos \theta}{13 - 12 \cos 2\theta} d\theta$$

$$9. \int_0^{2\pi} \frac{d\theta}{1 - 2k \sin \theta + k^2} \quad (k^2 < 1)$$

$$10. \int_0^{2\pi} \frac{\sin^2 \theta}{5 - 4 \cos \theta} d\theta$$

$$11. \int_0^{2\pi} \frac{\cos^2 \theta}{26 - 10 \cos 2\theta} d\theta$$

$$12. \int_0^{2\pi} \frac{\cos^2 3\theta}{5 - 4 \cos 2\theta} d\theta$$

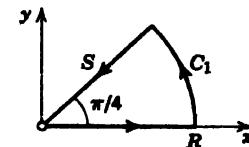


Fig. 350. Path in Example 5

13. Derive (10) from (9).

Evaluate the following integrals.

14. $\int_{-\infty}^{\infty} \frac{dx}{1+x^8}$

15. $\int_{-\infty}^{\infty} \frac{dx}{1+x^6}$

16. $\int_{-\infty}^{\infty} \frac{dx}{x^4 + 16}$

17. $\int_{-\infty}^{\infty} \frac{dx}{(1+x^8)^3}$

18. $\int_{-\infty}^{\infty} \frac{x}{(4+x^8)^2} dx$

19. $\int_{-\infty}^{\infty} \frac{x^3}{1+x^8} dx$

20. $\int_0^{\infty} \frac{1+x^3}{1+x^4} dx$

21. $\int_{-\infty}^{\infty} \frac{dx}{(x^8+1)(x^8+9)}$

22. $\int_{-\infty}^{\infty} \frac{dx}{(x^8+1)(x^8+4)^2}$

23. $\int_{-\infty}^{\infty} \frac{x}{(x^8-2x+2)^2} dx$

24. Solve Prob. 14 by elementary methods.

25. Obtain the answers to Probs. 18 and 19 without calculation.

Evaluate the following integrals.

26. $\int_{-\infty}^{\infty} \frac{\cos 2x}{(x^8+4)^2} dx$

27. $\int_{-\infty}^{\infty} \frac{\cos x}{(x^8+1)^2} dx$

28. $\int_{-\infty}^{\infty} \frac{\sin 3x}{1+x^4} dx$

29. $\int_0^{\infty} e^{-z^8} \cos 2xz dx$. Hint: integrate e^{-z^8} around the boundary of the rectangle with vertices at a , $-a$, $a + ib$, $-a + ib$, let $a \rightarrow \infty$, and use (16).

30. pr. v. $\int_{-\infty}^{\infty} \frac{dx}{(x+1)(x^8+2)}$. Hint: integrate $f(z) = 1/[(z+1)(z^8+2)]$ along the contour in Fig. 351.

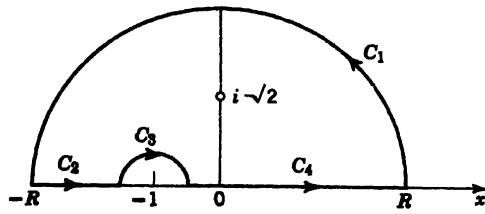


Fig. 351. Problem 30.

Chapter 16

Complex Analytic Functions and Potential Theory

Laplace's equation $\nabla^2 u = 0$ is one of the most important partial differential equations in engineering mathematics, because it occurs in connection with gravitational fields (Sec. 5.16), electrostatic fields (Sec. 9.11), steady-state heat conduction (Sec. 11.5), incompressible fluid flow, etc. The theory of the solutions of this equation is called *potential theory*.

In the "two-dimensional case" when u depends only on two Cartesian coordinates x and y , Laplace's equation becomes

$$\nabla^2 u = u_{xx} + u_{yy} = 0.$$

We know that then its solutions are closely related to complex analytic functions (cf. Sec. 10.3).¹ We shall now consider this connection and its consequences in more detail and illustrate it by practical examples taken from hydrodynamics and electrostatics. In Sec. 16.4 we shall see that results about analytic functions can be used for characterizing various general properties of harmonic functions. Finally, we shall derive an important general formula for the solution of boundary value problems involving Laplace's equation in a circular disk.

Prerequisites for this chapter: Chap. 10–13.

References: Appendix 1, Part G.

Answers to problems: Appendix 2.

16.1 *Electrostatic Fields*

The electrical force of attraction or repulsion between charged particles is governed by Coulomb's law. This force is the gradient of a function u , called the *electrostatic potential*, and at any points free of charges u is a solution of Laplace's equation

$$\nabla^2 u = 0.$$

¹ No such close relation exists in the three-dimensional case.

Cf. Sec. 9.11. The surfaces $u = \text{const}$ are called **equipotential surfaces**. At each point P the gradient of u is perpendicular to the surface $u = \text{const}$ through P , that is, the electrical force has the direction perpendicular to the equipotential surface.

Example 1. Find the potential of the field between two parallel conducting plates extending to infinity (Fig. 352), which are kept at potentials U_1 and U_2 , respectively. From the shape of the plates it follows that u depends only on x , and Laplace's equation becomes $u'' = 0$. By integrating twice, $u = ax + b$ where the constants a and b are determined by the given boundary values of u on the plates. For example, if the plates correspond to $x = -1$ and $x = 1$, the solution is

$$u(x) = \frac{1}{2}(U_2 - U_1)x + \frac{1}{2}(U_2 + U_1).$$

The equipotential surfaces are parallel planes.

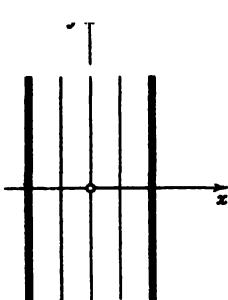


Fig. 352. Potential in Example 1.

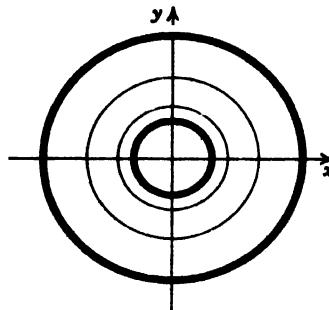


Fig. 353. Potential in Example 2.

Example 2. Find the potential between two coaxial conducting cylinders which extend to infinity on both sides (Fig. 353) and are kept at potentials U_1 and U_2 , respectively. Here u depends only on $r = \sqrt{x^2 + y^2}$, for reasons of symmetry, and Laplace's equation becomes

$$ru'' + u' = 0$$

[cf. (4) in Sec. 9.9].

By separating variables and integrating we obtain

$$u'' = -\frac{1}{r}, \quad \ln u' = -\ln r + a, \quad u' = \frac{a}{r}, \quad : a \ln r + b,$$

and a and b are determined by the given values of u on the cylinders. Although no infinitely extended conductors exist, the field in our idealized conductor will approximate the field in a long finite conductor in that part which is far away from the ends of the two cylinders.

If the potential u depends only on two Cartesian coordinates x and y , then Laplace's equation becomes

$$(1) \quad \nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

The equipotential surfaces $u = \text{const}$ appear as *equipotential lines* in the xy -plane.

We make the general assumption that $u(x, y)$ is *harmonic*, that is, its second partial derivatives are continuous. Then, if $v(x, y)$ is a conjugate

harmonic function of $u(x, y)$ (cf. Sec. 10.3), the function

$$F(z) = u(x, y) + iv(x, y)$$

is an analytic function of $z = x + iy$. This function is called the **complex potential** corresponding to the real potential u . Remember that for a given u the conjugate is uniquely determined, except for an additive real constant.

Since the lines $v = \text{const}$ intersect the equipotential lines $u = \text{const}$ at right angles [except at points where $F'(z) = 0$], they have the direction of the electrical force and, therefore, are called **lines of force**.

Example 3. In Ex. 1, a conjugate is $v = ay$. The complex potential is

$$F(z) = az + b = ax + b + iay,$$

and the lines of force are straight lines parallel to the x -axis.

Example 4. In Ex. 2 we have

$$u = a \ln r + b = a \ln |z| + b.$$

A conjugate is $v = a \arg z$. The complex potential is $F(z) = a \ln z + b$, and the lines of force are straight lines through the origin. The function $F(z)$ may also be interpreted as the complex potential of a source line whose trace with the xy -plane is the origin.

Often more complicated potentials can be obtained by superposition. This may be illustrated by

Example 5. Determine the complex potential of a pair of oppositely charged source lines of the same strength at the points $z = x_1$ and $z = x_2$. From Exs. 2 and 4 it follows that the potential of each of the source lines is

$$u_1 = -c \ln |z - x_1| \quad \text{and} \quad u_2 = c \ln |z - x_2|,$$

respectively. These are the real parts of the complex potentials

$$F_1(z) = -c \ln (z - x_1) \quad \text{and} \quad F_2(z) = c \ln (z - x_2).$$

Hence, the complex potential of the combination of the two source lines is

$$(2) \quad F(z) = F_1(z) + F_2(z) = c \ln \frac{z - x_2}{z - x_1}.$$

The equipotential lines are the circles

$$u = \operatorname{Re} F(z) = c \ln \frac{|z - x_2|}{|z - x_1|} = \text{const},$$

and the lines of force are the curves

$$v = \operatorname{Im} F(z) = c \arg \frac{z - x_2}{z - x_1} = c[\arg(z - x_2) - \arg(z - x_1)] = \text{const},$$

that is,

$$v = c(\theta_2 - \theta_1) = \text{const}$$

(cf. Fig. 354). Now, $|\theta_2 - \theta_1|$ is the angle between the line segments from z to x_1 and x_2 . The lines of force, therefore, are curves along each of which the line segment x_1x_2 appears

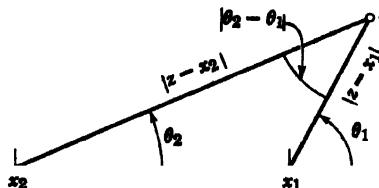


Fig. 354. Example 5.

under a constant angle; these curves are the totality of circular arcs over x_1x_2 , as is well known from elementary geometry. The function (2) may also be interpreted as the complex potential of a capacitor consisting of two circular cylinders whose axes are parallel but do not coincide (Fig. 355).

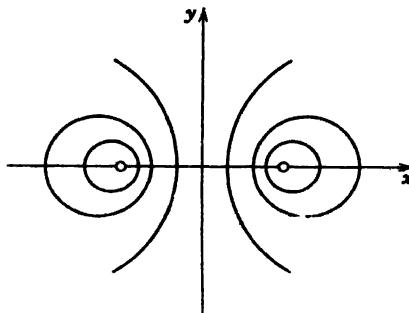


Fig. 355. Potential field of two oppositely charged source lines of the same strength.

Problems

Find the potential u between two infinite coaxial cylinders of radii r_1 and r_2 ($>r_1$) which are kept on the potentials U_1 and U_2 , respectively, where

1. $r_1 = 1, r_2 = 5, U_1 = 0, U_2 = 100 \text{ V}$
2. $r_1 = 0.5, r_2 = 2, U_1 = -110 \text{ V}, U_2 = 110 \text{ V}$
3. $r_1 = 2, r_2 = 20, U_1 = 100 \text{ V}, U_2 = 200 \text{ V}$
4. $r_1 = 3, r_2 = 6, U_1 = 100 \text{ V}, U_2 = 50 \text{ V}$

5. Find the equipotential lines of the complex potential $F(z) = 1/z$ and show these lines graphically.
6. Find the potential of two oppositely charged source lines at the points $z = a$ and $z = -a$. Show the equipotential lines graphically.
7. Find the potential of two source lines at $z = a$ and $z = -a$ having the same charge.
8. Show that $F(z) = \cos^{-1} z$ may be interpreted as the complex potential of the configurations in Figs. 356–358.

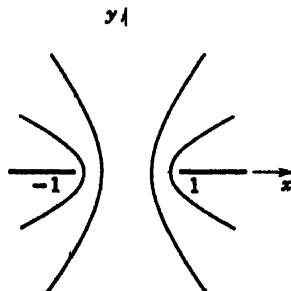


Fig. 356. Problem 8.

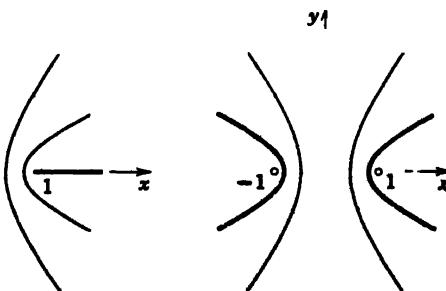


Fig. 357. Problem 8.

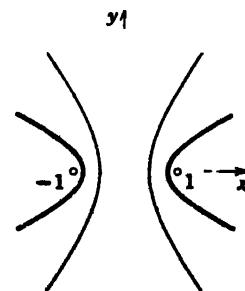


Fig. 358. Problem 8.

9. Show that $F(z) = \cosh^{-1} z$ may be interpreted as the complex potential between two confocal elliptic cylinders.
10. Find the potential u between the infinite cylinders in Fig. 359, if on the left cylinder, $u = -1$ and on the right, $u = 1$. Hint: use the potential in Prob. 6.

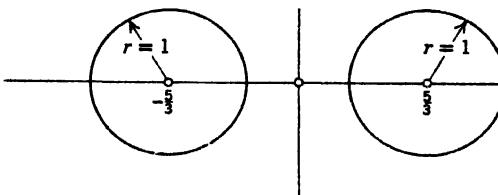


Fig. 359. Problem 10.

16.2 Two-Dimensional Fluid Flow

Harmonic functions play an important role in hydrodynamics. To illustrate this, let us consider the two-dimensional steady motion of a nonviscous fluid. Here "two-dimensional" means that the motion of the fluid is the same in all planes parallel to the xy -plane, the velocity being parallel to that plane. It then suffices to consider the motion of the fluid in the xy -plane. "Steady" means that the velocity is independent of the time.

At any point (x, y) the flow has a certain velocity which is determined by its magnitude and its direction and is thus a vector. Since in the complex plane any number a represents a vector (the vector from the origin to the point corresponding to a), we may represent the velocity of the flow by a complex variable, say,

$$(1) \quad V = V_1 + iV_2.$$

Then V_1 and V_2 are the components of the velocity in the x and y directions, and V is tangential to the paths of the moving particles of the fluid. Such a path is called a streamline of the motion. Cf. Fig. 360.

Consider now an arbitrary given smooth curve C . Let s denote the arc length of C , and let the real variable V_t be the component of the velocity V

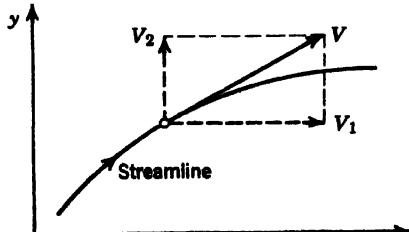
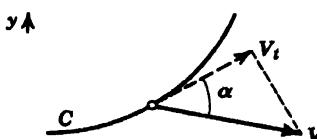


Fig. 360. Velocity.

Fig. 361. Tangential component of the velocity with respect to a curve C .

tangent to C (Fig. 361). Then the value of the line integral

$$(2) \quad \int_C V_t ds$$

taken along C in the sense of increasing values of s is called the **circulation** of the fluid along C . Dividing the circulation by the length of C , we obtain the mean velocity² of the flow along the curve C . Now

$$\cdot \quad V_t = |V| \cos \alpha \quad (\text{cf. Fig. 361}).$$

Consequently, V_t is the scalar product (Sec. 5.4) of V and the unit tangent vector (cf. Sec. 11.2)

$$\frac{dz}{ds} = \frac{dx}{ds} + i \frac{dy}{ds}$$

to C , where $z(s) = x(s) + iy(s)$ is a representation of C . Therefore, the product $V_t ds$ may be written

$$\cdot \quad V_t ds = V \cdot dz = V_1 dx + V_2 dy \quad (dz = dx + i dy).$$

Now let C be a *closed* curve, namely, the boundary curve of a simply connected domain D . Then, if V has continuous partial derivatives in a region containing D in its interior, it follows from Green's theorem (Sec. 6.4) that the circulation around C can be represented by a double integral,

$$(3) \quad \int_C (V_1 dx + V_2 dy) = \iint_D \left(\frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} \right) dx dy.$$

The function in the integral on the right permits a simple physical interpretation as follows. Let C be a circle of radius r . The circulation divided by $2\pi r$ then represents the mean velocity of the fluid along C , and the mean angular velocity ω_0 of the fluid about the axis of the circle is obtained by dividing the mean velocity by the radius r ; thus

$$\omega_0 = \frac{1}{\pi r^2} \iint_D \frac{1}{2} \left(\frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} \right) dx dy.$$

The expression on the right is the mean value³ of the function

$$(4) \quad \omega = \frac{1}{2} \left(\frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} \right)$$

² Definitions: $\frac{1}{b-a} \int_a^b f(x) dx$ = mean value of f on the interval $a \leq x \leq b$.

$\frac{1}{l} \int_C f(s) ds$ = mean value of f on C . (l = length of C .)

$\frac{1}{A} \iint_D f(x, y) dx dy$ = mean value of f on D (A = area of D .)

³ Cf. the previous footnote.

on the disk D bounded by the circle C . The function ω is called the **rotation**, and 2ω is called the **vorticity** of the motion. If $r \rightarrow 0$, the limit of that expression is the value of ω at the center of C . Hence, $\omega(x, y)$ is the limiting angular velocity of a circular element of the fluid as the circle shrinks to the point (x, y) . Roughly speaking, if a spherical element of the fluid were suddenly solidified and the surrounding fluid simultaneously annihilated, the element would rotate with the angular velocity ω (cf. also Sec. 5.19).

We consider only *irrotational flows*, that is, flows for which ω is zero everywhere in the region D of the flow, that is,

$$(5) \quad \frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} = 0,$$

the existence and continuity of the derivatives being assumed.

We further assume that the fluid is *incompressible*. Then

$$(6) \quad \frac{\partial V_1}{\partial x} + \frac{\partial V_2}{\partial y} = 0 \quad [\text{cf. (9), Sec. 5.18}]$$

in every region which is free of **sources** or **sinks**, that is, points at which fluid is produced or disappears.

If D is a simply connected region and the flow is irrotational, it follows from Theorem 3 in Sec. 6.12 that the line integral

$$(7) \quad \int_C (V_1 dx + V_2 dy)$$

is independent of path in D . If we integrate from a fixed point (a, b) in D to a variable point (x, y) in D , the integral thus becomes a function of the point (x, y) , say, $\Phi(x, y)$:

$$(8) \quad \Phi(x, y) = \int_{(a, b)}^{(x, y)} (V_1 dx + V_2 dy).$$

The function $\Phi(x, y)$ is called the **velocity potential**⁴ of the motion. Since the integral is independent of path, $V_1 dx + V_2 dy$ is an exact differential (Sec. 6.12), namely, the differential of the function $\Phi(x, y)$, that is,

$$(9) \quad V_1 dx + V_2 dy = \frac{\partial \Phi}{\partial x} dx + \frac{\partial \Phi}{\partial y} dy.$$

It follows that

$$(10) \quad V_1 = \frac{\partial \Phi}{\partial x}, \quad V_2 = \frac{\partial \Phi}{\partial y},$$

and we see that the velocity vector is the gradient of $\Phi(x, y)$:

$$(11) \quad V = V_1 + iV_2 = \frac{\partial \Phi}{\partial x} + i \frac{\partial \Phi}{\partial y} \quad (\text{cf. Sec. 5.16}).$$

⁴ Some authors use $-\Phi$ (instead of Φ) as the velocity potential.

The curves $\Phi(x, y) = \text{const}$ are called **equipotential lines**. Since V is the gradient of Φ , at each point V is perpendicular to the equipotential line through that point (provided $V \neq 0$).

Furthermore, by inserting (10) into (6) we see that Φ satisfies Laplace's equation

$$\nabla^2 \Phi = \frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} = 0.$$

Let $\Psi(x, y)$ be a conjugate harmonic function of $\Phi(x, y)$. Then at each point the curves

$$\Psi(x, y) = \text{const}$$

are perpendicular⁵ to the equipotential lines $\Phi(x, y) = \text{const}$. Hence, their tangents have the direction of the velocity of the fluid, and *the curves $\Psi(x, y) = \text{const}$ are, therefore, the streamlines of the flow*. The function $\Psi(x, y)$ is called the **stream function** of this flow.

Assuming that both Φ and Ψ have continuous second partial derivatives, the complex function

$$(12) \quad F(z) = \Phi(x, y) + i\Psi(x, y)$$

is analytic in the region of the flow. This function is called the **complex potential** of the flow. Working with the complex potential is simpler than working with Φ and Ψ separately.

The velocity of the flow can be obtained by differentiating (12) and using the Cauchy-Riemann equations; we find

$$F'(z) = \frac{\partial \Phi}{\partial x} + i \frac{\partial \Psi}{\partial x} = \frac{\partial \Phi}{\partial x} - i \frac{\partial \Phi}{\partial y} = V_1 - iV_2.$$

From this it follows that

$$(13) \quad V = V_1 + iV_2 = \bar{F}'(z).$$

In this way *two-dimensional irrotational steady flows of incompressible fluids may be described in terms of analytic functions*, and the methods of complex analysis, such as conformal mapping, may be employed.

In connection with boundary value problems the stream function Ψ is important, because a boundary across which fluid cannot flow is a streamline. Under conformal mapping, a streamline transforms into a streamline in the image plane. Another possibility for obtaining and investigating complicated flows is the *superposition of simple flows*. The sum $F = F_1 + F_2$ of the complex potentials F_1, F_2 of two flows is the complex potential of the flow which is obtained by vector addition of the velocity vectors of the two flows. Clearly, since Laplace's equation is linear and homogeneous, the sum of two harmonic functions is a harmonic function.

Special examples will be considered in the next section.

⁵ Except for points where $F'(z) = 0$ [cf. (12)].

Note that while in electrostatics the given boundaries (conducting plates) are equipotential lines, in hydrodynamics the boundaries are streamlines and thus are orthogonal to the equipotential lines.

16.3 Special Complex Potentials

The following examples may illustrate that even the simplest analytic functions correspond to interesting flow patterns.

Note that some functions considered in Sec. 16.1 now reappear, but have an entirely different physical interpretation. The problem of uniqueness is discussed in Ref. [C5] and will not be considered here.

Example 1 (Parallel flow). The complex potential

$$(1) \quad F(z) = Kz \quad (K \text{ positive real})$$

describes a uniform flow to the right whose equipotential lines are

$$\operatorname{Re} F(z) = Kx = \text{const}$$

and whose streamlines are

$$\operatorname{Im} F(z) = Ky = \text{const},$$

that is, horizontal straight lines. From (13), Sec. 16.2, we obtain

$$V = V_1 + iV_2 = K, \quad \text{that is,} \quad V_1 = K, \quad V_2 = 0;$$

the velocity vector is constant and parallel to the x -axis, and the flow can be interpreted as a uniform flow between two parallel lines (between two parallel planes in three-dimensional space). Cf. Fig. 362.

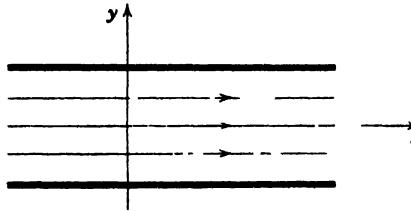


Fig. 362. Parallel flow.

Example 2 (Flow around a corner). The complex potential

$$(2) \quad F(z) = z^2 = x^2 - y^2 + 2ixy$$

describes a flow whose equipotential lines are the hyperbolas

$$\Phi = x^2 - y^2 = \text{const}$$

and whose streamlines are the hyperbolas

$$\Psi = 2xy = \text{const.}$$

From (13) in the last section we obtain the velocity vector

$$V = 2z = 2(x - iy), \quad \text{that is,} \quad V_1 = 2x, \quad V_2 = -2y.$$

The speed (magnitude of the velocity) is

$$|V| = \sqrt{V_1^2 + V_2^2} = 2\sqrt{\omega^2 + y^2}.$$

The flow may be interpreted as the flow in a channel bounded by the positive coordinate axes and a hyperbola, say, $xy = 1$ (Fig. 363). We note that the speed along a streamline S has a minimum at the point P where the cross section of the channel is large.

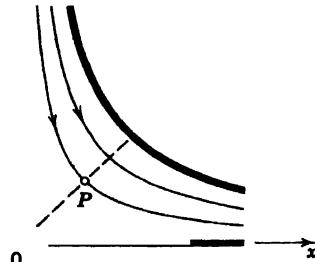


Fig. 363. Flow around a corner.

Example 3 (Source and sink). The complex potential

$$(3) \quad F(z) = \frac{c}{2\pi} \ln z \quad (c \text{ positive real})$$

corresponds to a point source at $z = 0$ (this means a source line $x = 0, y = 0$ in space). The constant c is called the strength or discharge of the source. The equipotential lines

$$\Phi = \operatorname{Re} F(z) = \frac{c}{2\pi} \ln r = \operatorname{const} \quad (r = \sqrt{x^2 + y^2})$$

are concentric circles, and the streamlines

$$\psi = \operatorname{Im} F(z) = \frac{c\theta}{2\pi} = \operatorname{const} \quad (\theta = \arg z)$$

are straight lines through the origin. Since

$$V = \overline{F'(z)} = \frac{c}{2\pi z} = \frac{c}{2\pi r^2}(x + iy),$$

we see that the flow is directed radially outward. Cf. Fig. 364.

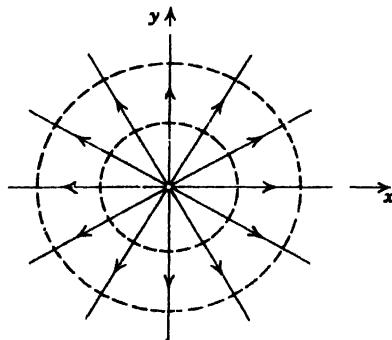


Fig. 364. Point source.

If c is negative real, then the flow described by (3) is said to have a sink at $z = 0$; the flow is directed radially inward, and fluid disappears at the singular point $z = 0$ of the complex potential.

Example 4 (Vortex line). If in (3) the constant c is purely imaginary, say, $c = -iK$ where $K > 0$, then

$$(4) \quad F(z) = -\frac{iK}{2\pi} \ln z.$$

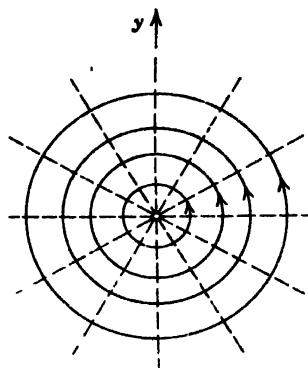


Fig. 365. Vortex flow.

The equipotential lines are the straight lines (cf. Fig. 365)

$$\Phi = \operatorname{Re} F(z) = \frac{K}{2\pi} \theta = \text{const.}$$

and the streamlines are the concentric circles

$$\Psi = \operatorname{Im} F(z) = -\frac{K}{2\pi} \ln r = \text{const.}$$

Hence, the flow circles around the origin in the counterclockwise sense. The point $z = 0$ is a **vortex**. The potential increases by the value K each time we travel around the vortex.

Example 5 (Flow around a cylinder). We shall now consider the complex potential

$$F(z) = z + \frac{1}{z}.$$

Setting $z = re^{i\theta}$, we obtain

$$\begin{aligned} F(z) &= re^{i\theta} + \frac{1}{r} e^{-i\theta} \\ &= \left(r + \frac{1}{r}\right) \cos \theta + i \left(r - \frac{1}{r}\right) \sin \theta. \end{aligned}$$

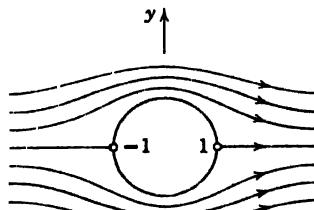


Fig. 366. Flow around a cylinder.

Consequently, the streamlines are

$$\Psi(x, y) = \left(r - \frac{1}{r}\right) \sin \theta = \text{const.}$$

The particular streamline

$$\left(r - \frac{1}{r}\right) \sin \theta = 0$$

consists of the x -axis ($\theta = 0$ and $\theta = \pi$) and the unit circle $r = 1$. For large $|z|$ we have $F(z) \approx z$ and the flow is then nearly uniform and parallel. It may be interpreted as the flow around a long circular cylinder of unit radius (Fig. 366). The velocity is

$$V = \overline{F'(z)} = 1 - \frac{1}{z^2}.$$

At $z = -1$ and $z = 1$ we have $V = 0$. Such points are called **stagnation points** of the flow.

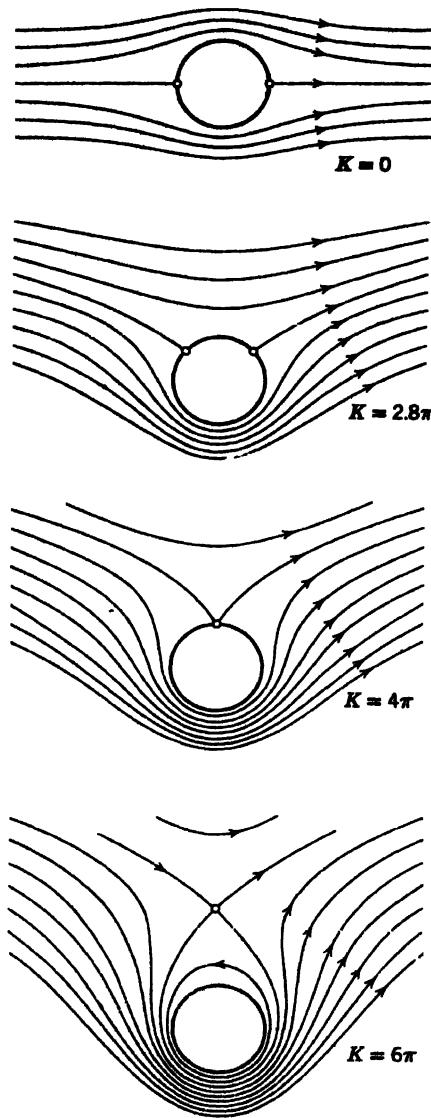


Fig. 367. Flow around a cylinder with circulation.

Example 6 (Flow with circulation around a cylinder). This flow can be obtained by superposition of the flows in Exs. 4 and 5. In both flows the cylinder wall $|z| = 1$ is a streamline. By adding the two corresponding complex potentials we obtain

$$(5) \quad F(z) = z + \frac{1}{z} - \frac{iK}{2\pi} \ln z.$$

The streamlines are

$$\Psi = \operatorname{Im} F(z) = y - \frac{\pi}{x^2 + y^2} - \frac{iK}{2\pi} \ln \sqrt{x^2 + y^2} = \text{const.}$$

Depending on the magnitude of K , we may distinguish between three types of flow having either two or one or no stagnation points on the cylinder wall. The speed is

$$|V| = |\overline{F'(z)}| = |F'(z)| = \left| \left(1 - \frac{1}{z^2} \right) - \frac{iK}{2\pi z} \right|.$$

We first note that $|V| \rightarrow 1$ as $|z| \rightarrow \infty$; that is, for points at a great distance from the cylinder the flow is nearly parallel and uniform. The stagnation points are the solution of the equation $V = 0$, that is,

$$(6) \quad z^2 - \frac{iK}{2\pi} z - i = 0.$$

We obtain

$$z = \frac{iK}{4\pi} \pm \sqrt{\frac{-K^2}{16\pi^2} + 1}.$$

If $K = 0$ (no circulation), then $z = \pm 1$, as in Ex. 5. As K increases from 0 to 4π , the stagnation points move from $z = \pm 1$ up on the unit circle until they unite at $z = i$. The value $K = 4\pi$ corresponds to a double root of the equation (6). If $K > 4\pi$, the roots of (6) become imaginary, so that one of the stagnation points lies on the imaginary axis in the field of flow while the other one lies inside the cylinder, thus losing its physical meaning (Fig. 367).

Problems

1. Graph some of the equipotential lines in Example 5.
2. Obtain the flow in Ex. 2 from that in Ex. 1 by a conformal mapping of the first quadrant onto the upper half-plane.
3. In Ex. 5 determine the velocity at the points $\pm 1, \pm 2, \pm 3, \pm 4$, and on the wall of the cylinder.

Consider the flow corresponding to the given complex potential $F(z)$. Show the streamlines and equipotential lines graphically. Find the velocity vector, and determine all points at which that vector is parallel to the x -axis.

- | | | |
|---------------------|------------------|----------------------|
| 4. $F(z) = iz$ | 5. $F(z) = -ikz$ | 6. $F(z) = (1 + i)z$ |
| 7. $F(z) = z^2 + z$ | 8. $F(z) = iz^2$ | 9. $F(z) = z^3$ |
| 10. $F(z) = iz^3$ | 11. $F(z) = z^4$ | 12. $F(z) = i \ln z$ |
-
13. Find the complex potential of a flow which has a point source of strength 1 at $z = -a$.
 14. Find the complex potential of a flow which has a point sink of strength 1 at $z = a$.
 15. Show that vector addition of the velocity vectors of two flows leads to a flow whose complex potential is obtained by adding those of the two flows.
 16. Add the potentials in Probs. 13 and 14 and show the streamlines graphically.
 17. Find the streamlines of the flow corresponding to $F(z) = 1/z$. Show that for small $|a|$ the streamlines in Prob. 16 look similar to those in the present problem.
 18. Show that $F(z) = \cosh^{-1} z$ corresponds to a flow whose streamlines are confocal hyperbolae with foci at $z = \pm 1$, and the flow may be interpreted as a flow through an aperture (cf. Fig. 368 on the next page).
 19. Show that $F(z) = \cos^{-1} z$ can be interpreted as the complex potential of a flow circulating around an elliptic cylinder or around a plate (the straight segment from

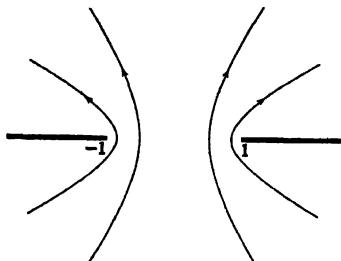


Fig. 368. Flow through an aperture.

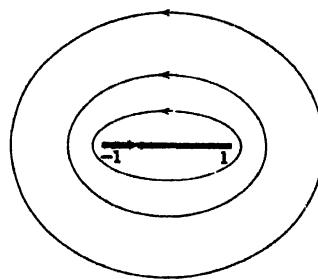


Fig. 369. Flow around a plate.

$z = -1$ to $z = 1$). Show that the streamlines are confocal ellipses with foci at $z = \pm 1$. (Cf. Fig. 369.)

20. Find a representation of the curves along which the speed (magnitude of the velocity) of a flow corresponding to a complex potential $F(z)$ is constant.

16.4 General Properties of Harmonic Functions

In this section we want to show how results about complex analytic functions can be used for deriving general properties of harmonic functions.

If $u(x, y)$ is any given harmonic function, we may determine a conjugate harmonic function $v(x, y)$ by the use of the Cauchy-Riemann equations, and $f(z) = u(x, y) + iv(x, y)$ is then an analytic function (cf. Sec. 10.3). This shows that any harmonic function can be regarded as the real part of a suitable analytic function. Since analytic functions have derivatives of all orders, we immediately obtain the following result.

Theorem 1. *A function $u(x, y)$ which is harmonic in a domain D has partial derivatives of all orders in D .*

If $f(z)$ is analytic in a simply connected domain D , then, by Cauchy's integral formula (Sec. 12.5),

$$(1^*) \quad f(z_0) = \frac{1}{2\pi i} \int_C \frac{f(z)}{z - z_0} dz$$

where C is a simple closed path in D and z_0 lies inside C . Choosing for C a circle

$$z = z_0 + re^{i\phi}$$

in D , we have

$$z - z_0 = re^{i\phi}, \quad dz = ire^{i\phi} d\phi$$

and the integral formula becomes

$$\cdot (1) \quad f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{i\phi}) d\phi.$$

The right side is the mean value of $f(z)$ on the circle (= value of the integral divided by the length of the interval of integration). This proves

Theorem 2 (Mean value property). *Let $f(z)$ be analytic in a simply connected domain D . Then the value of $f(z)$ at a point z_0 in D is equal to the mean value of $f(z)$ on any circle in D with center at z_0 .*

Another important property of analytic functions is the following one.

Theorem 3 (Maximum modulus theorem). *If $f(z)$ is analytic and nonconstant in a bounded region D and on the boundary of D , then its absolute value $|f(z)|$ cannot have a maximum at an interior point of D . Consequently, the maximum of $|f(z)|$ is taken on the boundary of D . If $f(z) \neq 0$ in D , the same is true with respect to the minimum of $|f(z)|$.*

Proof. We show that the assumption that $|f(z)|$ has a maximum at an interior point of D leads to a contradiction. Let z_0 be that point, and let $|f(z_0)| = M$ be that maximum. Since $f(z)$ is not constant, $|f(z)|$ is not constant. Consequently, we can find a circle C of radius r with center at z_0 such that the interior of C is in D and $|f(z)|$ is smaller than M at some point P of C . Since $|f(z)|$ is continuous, it will be smaller than M on an arc C_1 of C which contains P , say, $|f(z)| \leq M - \epsilon$ ($\epsilon > 0$) for all z on C_1 (Fig. 370). If C_1 has the length l_1 , the complementary arc C_2 of C has the length $2\pi r - l_1$. By applying (4), Sec. 12.2, to (1*), this section, and noting that $|z - z_0| = r$, we thus obtain

$$\begin{aligned} M = |f(z_0)| &\leq \frac{1}{2\pi} \left| \int_{C_1} \frac{f(z)}{z - z_0} dz \right| + \frac{1}{2\pi} \left| \int_{C_2} \frac{f(z)}{z - z_0} dz \right| \\ &\leq \frac{1}{2\pi} (M - \epsilon) \frac{1}{r} l_1 + \frac{1}{2\pi} M \frac{1}{r} (2\pi r - l_1) = M - \frac{\epsilon l_1}{2\pi r} < M, \end{aligned}$$

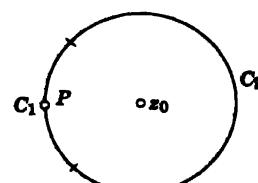


Fig. 370. Proof of Theorem 3.

that is, $M < M$, which is impossible. Hence, our assumption is false and the first statement of the theorem is proved.

We prove the last statement. If $f(z) \neq 0$ in D , then $1/f(z)$ is analytic in D . From the statement already proved it follows that the maximum of $1/|f(z)|$ lies on the boundary of D . But this maximum corresponds to a minimum of $|f(z)|$. This completes the proof.

From these theorems we shall now derive corresponding results about harmonic functions.

Theorem 4. *Let D be a simply connected bounded domain and let C be its boundary curve. Then if $u(x, y)$ is harmonic in a region containing D in its interior, it has the following properties.*

I. The value of $u(x, y)$ at an interior point (x_0, y_0) of D is equal to the mean value of $u(x, y)$ on any circle in D with center at (x_0, y_0) .

II. The value of $u(x, y)$ at the point (x_0, y_0) is equal to the mean value of $u(x, y)$ in any circular disk in D with center at (x_0, y_0) . [Cf. footnote 2 in Sec. 16.2.]

III. (Maximum principle) If $u(x, y)$ is nonconstant, it has neither a maximum nor a minimum in D . Consequently, the maximum and the minimum are taken on the boundary of D .

IV. If $u(x, y)$ is constant on C , then $u(x, y)$ is a constant.

V. If $v(x, y)$ is harmonic in D and on C and if $v(x, y) = u(x, y)$ on C , then $v(x, y) = u(x, y)$ everywhere in D .

Proof. Statement I follows from (1) by taking the real parts on both sides, that is,

$$u(x_0, y_0) = \operatorname{Re} f(x_0 + iy_0) = \frac{1}{2\pi} \int_0^{2\pi} u(x_0 + r \cos \phi, y_0 + r \sin \phi) d\phi.$$

If we multiply both sides by r and integrate over r from 0 to r_0 where r_0 is the radius of a circular disk in D with center at (x_0, y_0) , then we obtain on the left-hand side $\frac{1}{2}r_0^2 u(x_0, y_0)$ and therefore

$$u(x_0, y_0) = \frac{1}{\pi r_0^2} \int_0^{r_0} \int_0^{2\pi} u(x_0 + r \cos \phi, y_0 + r \sin \phi) r d\phi dr.$$

This proves the second statement.

We prove statement III. Let $v(x, y)$ be a conjugate harmonic function of $u(x, y)$ in D . Then $f(z) = u(x, y) + iv(x, y)$ is analytic in D , and so is

$$F(z) = e^{f(z)}.$$

The absolute value is

$$|F(z)| = e^{\operatorname{Re} f(z)} = e^{u(x, y)}.$$

From Theorem 3 it follows that $|F(z)|$ cannot have a maximum at an interior point of D . Since e^u is a monotone increasing function of the real variable u , statement III about the maximum of u follows. From this, the statement about the minimum follows by replacing u by $-u$.

If u is constant on C , say, $u = k$, then by III the maximum and the minimum of u are equal. From this, statement IV follows.

If u and v are harmonic in D and on C , then $u - v$ is also harmonic in D and on C , and by assumption, $u - v = 0$ everywhere on C . Hence, by IV, $u - v = 0$ everywhere in D , and statement V is proved. This completes the proof of Theorem 4.

The last statement of Theorem 4 is very important. It means that a harmonic function is uniquely determined in D by its values on the boundary of D . Usually, $u(x, y)$ is required to be harmonic in D and continuous on the boundary* of D . Under these circumstances the maximum principle

* That is, $\lim_{\substack{z \rightarrow z_0 \\ z \in D}} u(x, y) = u(x_0, y_0)$ where (x_0, y_0) is on the boundary and (x, y) is in D .

(Theorem 4, III) is still applicable. The problem of determining $u(x, y)$ when the boundary values are given is known as **Dirichlet's problem**. From Theorem 4, V we have

Theorem 5. *If for a given region and given boundary values the Dirichlet problem has a solution, the solution is unique.*

Problems

1. Verify Theorem 2 for $f(z) = (z + 2)^3$, $z_0 = 1$ and a circle of radius 1 with center at z_0 .

Verify Theorem 3 for:

2. $f(z) = z^2$ and D : $-2 \leq x \leq 2$, $-1 \leq y \leq 1$.
3. $f(z) = e^z$ and any bounded region R .
4. The function $f(x) = \cos x$ has a maximum at $x = 0$. Using Theorem 3, conclude that the modular surface of $f(z) = \cos z$ (cf. Sec. 10.6) cannot have a summit at $z = 0$.
5. If $f(z)$ is analytic (not a constant) in a simply connected domain D , and the curve given by $|f(z)| = c$ (c any fixed constant) lies in D and is closed, show that $f(z) = 0$ at a point in the interior of that curve. Give examples.

16.5 Poisson's Integral Formula

The Dirichlet problem for a circular disk can be solved by the use of the so-called Poisson⁷ formula which represents a harmonic function in terms of its values given on the boundary circle of the disk. We shall derive this formula, starting from Cauchy's integral formula

$$(1) \quad f(z) = \frac{1}{2\pi i} \int_C \frac{f(z^*)}{z^* - z} dz^*;$$

here C is the circle represented by

$$z^* = Re^{i\phi} \quad (0 \leq \phi \leq 2\pi)$$

and the function

$$f(z) = u(r, \theta) + iv(r, \theta) \quad (z = re^{i\theta})$$

is assumed to be analytic in a simply connected region containing C in its interior.

Since $dz^* = iRe^{i\phi} d\phi = iz^* d\phi$, we obtain from (1)

$$(2) \quad f(z) = \frac{1}{2\pi} \int_0^{2\pi} f(z^*) \frac{z^*}{z^* - z} d\phi \quad (z^* = Re^{i\phi}, z = re^{i\theta}).$$

⁷ SIMÉON DENIS POISSON (1781–1840), French mathematician and physicist.

On the other hand, if we consider a point Z outside C , say, the point $Z = z^* \bar{z}^*/\bar{z}$ (whose absolute value is $R^2/r > R$), then the integrand in (1) is analytic in the disk $|z| \leq R$ and the integral is zero:

$$0 = \frac{1}{2\pi i} \int_C \frac{f(z^*)}{z^* - Z} dz^* = \frac{1}{2\pi} \int_0^{2\pi} f(z^*) \frac{z^*}{z^* - Z} d\phi.$$

By inserting $Z = z^* \bar{z}^*/\bar{z}$ and simplifying the fraction we find

$$0 = \frac{1}{2\pi} \int_0^{2\pi} f(z^*) \frac{\bar{z}}{\bar{z} - \bar{z}^*} d\phi.$$

We subtract this from (2); then, since

$$(3) \quad \frac{z^*}{z^* - z} - \frac{\bar{z}}{\bar{z} - \bar{z}^*} = \frac{z^* \bar{z}^* - z \bar{z}}{(z^* - z)(\bar{z}^* - \bar{z})}$$

we obtain

$$(4) \quad f(z) = \frac{1}{2\pi} \int_0^{2\pi} f(z^*) \frac{z^* \bar{z}^* - z \bar{z}}{(z^* - z)(\bar{z}^* - \bar{z})} d\phi.$$

From the polar representations of z and z^* we see that the quotient in the integrand is equal to

$$(Re^{i\phi} - re^{i\theta})(Re^{-i\phi} - re^{-i\theta}) = \frac{R^2 - r^2}{R^2 - 2Rr \cos(\theta - \phi) + r^2}.$$

Hence, by taking the real parts on both sides of (4) we obtain Poisson's integral formula

$$(5) \quad u(r, \theta) = \frac{1}{2\pi} \int_0^{2\pi} u(R, \phi) \frac{R^2 - r^2}{R^2 - 2Rr \cos(\theta - \phi) + r^2} d\phi,$$

which represents the harmonic function u in the disk $|z| \leq R$ in terms of its values $u(R, \phi)$ on the circle which bounds the disk.

It is of practical interest to note that in (5) we may use for $u(R, \phi)$ any function which is merely piecewise continuous on the interval of integration. Formula (5) then defines a function $u(r, \theta)$ which is harmonic in the open disk $|z| < R$ and continuous on the circle $|z| = R$. On this circle this function is equal to $u(R, \phi)$, except at points where $u(R, \phi)$ is discontinuous. The proof can be found in Ref. [G1].

From (5) we may obtain an important series development of u in terms of simple harmonic functions. We remember that the quotient in the integrand of (5) was derived from (3), and it is not difficult to see that the right side of (3) is the real part of $(z^* + z)/(z^* - z)$. By the use of the geometric series we obtain

$$(6) \quad \frac{z^* + z}{z^* - z} = \frac{1 + (z/z^*)}{1 - (z/z^*)} = \left(1 + \frac{z}{z^*}\right) \sum_{n=0}^{\infty} \left(\frac{z}{z^*}\right)^n = 1 + 2 \sum_{n=1}^{\infty} \left(\frac{z}{z^*}\right)^n.$$

Since $z = re^{i\theta}$ and $z^* = Re^{i\phi}$, we have

$$(7) \quad \begin{aligned} \operatorname{Re} \left(\frac{z}{z^*} \right)^n &= \operatorname{Re} \left[\frac{r^n}{R^n} e^{in\theta} e^{-in\phi} \right] = \left(\frac{r}{R} \right)^n \cos(n\theta - n\phi) \\ &= \left(\frac{r}{R} \right)^n (\cos n\theta \cos n\phi + \sin n\theta \sin n\phi). \end{aligned}$$

From (6) and (7) we obtain

$$\operatorname{Re} \frac{z^* + z}{z^* - z} = 1 + 2 \sum_{n=1}^{\infty} \frac{r^n}{R^n} (\cos n\theta \cos n\phi + \sin n\theta \sin n\phi).$$

This expression is equal to the quotient in (5), as we have mentioned before, and by inserting the series into (5) and integrating term by term we find

$$(8) \quad u(r, \theta) = a_0 + \sum_{n=1}^{\infty} \left(\frac{r}{R} \right)^n (a_n \cos n\theta + b_n \sin n\theta)$$

where the coefficients are

$$(9) \quad \begin{aligned} a_0 &= \frac{1}{2\pi} \int_0^{2\pi} u(R, \phi) d\phi, \quad a_n = \frac{1}{\pi} \int_0^{2\pi} u(R, \phi) \cos n\phi d\phi, \\ b_n &= \frac{1}{\pi} \int_0^{2\pi} u(R, \phi) \sin n\phi d\phi, \quad n = 1, 2, \end{aligned}$$

the Fourier coefficients of $u(R, \phi)$. Note that for $r = R$ the series (8) becomes the Fourier series of $u(R, \phi)$, and therefore the representation (8) will be valid whenever $u(R, \phi)$ can be represented by a Fourier series.

Example 1. Find the potential $u(r, \theta)$ in the unit disk $r < 1$ having the boundary values (Fig. 371)

$$u(1, \phi) = \begin{cases} -\phi/\pi & \text{when } -\pi < \phi < 0 \\ \phi/\pi & \text{when } 0 < \phi < \pi. \end{cases}$$

Since $u(1, \phi)$ is even, $b_n = 0$, and from (9) we obtain $a_0 = 1/2$ and

$$a_n = \frac{1}{\pi} \left[- \int_{-\pi}^0 \frac{\phi}{\pi} \cos n\phi d\phi + \int_0^\pi \frac{\phi}{\pi} \cos n\phi d\phi \right] = \frac{-1}{n^2 \pi^2} (\cos n\pi - 1).$$

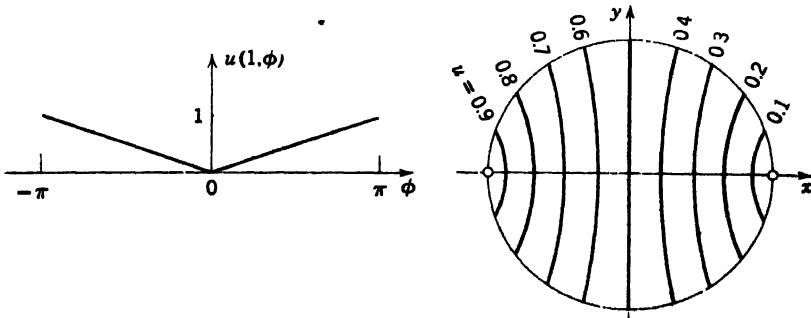


Fig. 371. Potential in Example 1.

Hence, $a_n = -4/n^2\pi^2$ when n is odd, $a_n = 0$ when $n = 2, 4, \dots$, and the potential is

$$u(r, \theta) = \frac{1}{2} - \frac{4}{\pi^2} \left[r \cos \theta + \frac{r^3}{3^2} \cos 3\theta + \frac{r^5}{5^2} \cos 5\theta + \dots \right].$$

Problems

1. Verify (3).

2. Show that each term in (8) is a harmonic function in the disk $r^2 < R^2$.

Using (8), find the potential $u(r, \theta)$ in the unit disk $r < 1$ having the given boundary values $u(1, \theta)$. Using the first few terms of the series, compute some values of u and sketch a figure of the equipotential lines.

3. $u(1, \theta) = \sin \theta$

4. $u(1, \theta) = 1 - \cos \theta$

5. $u(1, \theta) = \sin 3\theta$

6. $u(1, \theta) = \cos 2\theta - \cos 4\theta$

7. $u(1, \theta) = 4 \sin^3 \theta$

8. $u(1, \theta) = \theta$

9. $u(1, \theta) = 1$ if $0 < \theta < \pi$ and 0 otherwise

10. $u(1, \theta) = \theta$ if $-\pi/2 < \theta < \pi/2$, $u(1, \theta) = \pi - \theta$ if $\pi/2 < \theta < 3\pi/2$

11. $u(1, \theta) = -\pi/2$ if $-\pi < \theta < -\pi/2$, $u(1, \theta) = \theta$ if $-\pi/2 < \theta < \pi/2$,
 $u(1, \theta) = \pi/2$ if $\pi/2 < \theta < \pi$

12. $u(1, \theta) = 1$ if $0 < \theta < \pi/2$, $u(1, \theta) = -1$ if $\pi/2 < \theta < \pi$ and 0 otherwise

13. Using (9) in Sec. 14.2, show that the result of Prob. 12 may be written

$$u(r, \theta) = \frac{1}{\pi} \operatorname{Im} \ln \frac{(1+iz)(1+z^2)}{(1-iz)(1-z^2)}.$$

14. Show that the potential in Prob. 8 may be written $u(r, \theta) = 2 \operatorname{Im} \ln(1+z)$.

15. Using (8), show that the potential $u(r, \theta)$ in the unit disk $r < 1$ having the boundary values

$$u(1, \theta) = \begin{cases} -1 & \text{when } -\pi < \theta < 0 \\ 1 & \text{when } 0 < \theta < \pi \end{cases}$$

is given by the series

$$u(r, \theta) = \frac{4}{\pi} \left(r \sin \theta + \frac{r^3}{3} \sin 3\theta + \frac{r^5}{5} \sin 5\theta + \dots \right).$$

Compute some values of u by the use of the first few terms of this series and draw some of the equipotential lines (Fig. 372).

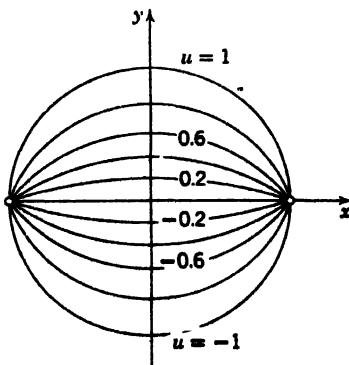


Fig. 372. Problem 15.

16. Using (9), Sec. 14.2, show that in Prob. 15

$$u(r, \theta) = \frac{2}{\pi} \operatorname{Im} \ln \frac{1+z}{1-z} = \frac{2}{\pi} [\arg(1+z) - \arg(1-z)].$$

17. Applying a familiar theorem of elementary geometry to the result of Prob. 16, show that the curves $u = \text{const}$ in Fig. 372 are circular arcs.

18. Show that

$$U = 1 + \frac{2}{\pi} \operatorname{Im} \ln \frac{w+1}{w-1} \quad (w = u+iv)$$

is harmonic in the upper half-plane $v > 0$ and has the values -1 for $-1 < u < 1$ and $+1$ on the remaining part of the u -axis

19. Show that the linear transformation which maps $w_1 = -1$, $w_2 = 0$, $w_3 = 1$ onto $z_1 = -1$, $z_2 = -i$, $z_3 = 1$, respectively, is

$$z = \frac{w-i}{-iw+1}.$$

Find the inverse $w = w(z)$, insert it into U in Prob. 18, and show that the resulting harmonic function is that in Prob. 16.

20. Derive Theorem 4 I, Sec. 16.4, from (5).

Chapter 17

Special Functions. Asymptotic Expansions

Two classes of nonelementary functions play an important role in engineering mathematics. The first class consists of functions arising in connection with linear differential equations; examples of such functions were considered in Chap. 3. The second class consists of functions defined by integrals which cannot be evaluated in terms of finitely many elementary functions. It includes the gamma function, beta function, error function, sine and cosine integrals, and Fresnel integrals, which will be considered in the present chapter.

We shall see that some of these functions (and other functions) have so-called *asymptotic expansions*. These are series which may not converge, but can nevertheless be used for computing numerical values of the functions for large values of the independent variable. We shall see that, roughly speaking, the greater the values of the independent variable are, the less work will be required in those computations. Important examples of asymptotic series will be considered in Sec. 17.3 and general properties of these series in Sec. 17.4.

Prerequisites for this chapter: elementary integral calculus.

References: Appendix 1, Part H.

Answers to problems: Appendix 2.

17.1 Gamma and Beta Functions

One of the most important nonelementary functions is the **gamma function**, written $\Gamma(\alpha)$ and defined by the integral

$$(1) \quad \Gamma(\alpha) = \int_0^{\infty} e^{-t} t^{\alpha-1} dt \quad (\alpha > 0).$$

The gamma function occurs in connection with various practical and theoretical considerations in physics, engineering, and mathematical statistics, and therefore the student should become familiar with its properties.

The integral representation (1) is meaningful only when α is positive.¹ To show this, write

$$(2) \quad \Gamma(\alpha) = \int_0^c e^{-t} t^{\alpha-1} dt + \int_c^\infty e^{-t} t^{\alpha-1} dt,$$

where c is any positive number. While the last integral on the right exists for all $c > 0$ and all α (as the reader may prove), the situation with respect to the first integral is different. By using the mean-value theorem for integrals we obtain

$$\int_0^c e^{-t} t^{\alpha-1} dt = e^{-b} \int_0^c t^{\alpha-1} dt = e^{-b} \frac{c^\alpha}{\alpha} \quad (\alpha \neq 0),$$

where b is some real number between 0 and c . If $\alpha > 0$, then $t^\alpha \rightarrow 0$ as $t \rightarrow 0$ so that the expression on the right and, therefore, also the integral on the left, exists. On the other hand, for $\alpha < 0$ we have $t^\alpha \rightarrow \infty$ as $t \rightarrow 0$, and that integral does not exist. For $\alpha = 0$ we obtain $e^{-b} \ln t$ which becomes infinite as $t \rightarrow 0$.

By integration by parts we find

$$\Gamma(\alpha + 1) = \int_0^\infty e^{-t} t^\alpha dt = -e^{-t} t^\alpha \Big|_0^\infty + \alpha \int_0^\infty e^{-t} t^{\alpha-1} dt \quad (\alpha > 0).$$

The first expression on the right is zero, and the integral on the right is $\Gamma(\alpha)$. We thus obtain the important *functional relation of the gamma function*,

$$(3) \quad \Gamma(\alpha + 1) = \alpha \Gamma(\alpha).$$

Let α be a positive integer, say, k . Then repeated application of (3) yields

$$\begin{aligned} \Gamma(k + 1) &= k \Gamma(k) = k(k - 1) \Gamma(k - 1) \\ &\quad \cdots \cdots \\ &= k(k - 1) \cdots 3 \cdot 2 \cdot 1 \cdot \Gamma(1). \end{aligned}$$

Now, for the remaining function $\Gamma(1)$ we obtain by integration

$$(4) \quad \Gamma(1) = \int_0^\infty e^{-t} dt = 1.$$

This proves the basic relation

$$(5) \quad \Gamma(k + 1) = k! \quad (k = 0, 1, \dots).$$

Hence, the gamma function can be regarded as a generalization of the elementary factorial function; in fact, this property was the historical reason for introducing the gamma function.²

¹ Or, more generally, if we consider also complex values, for those α whose real part is positive.

² Sometimes the notation $(x - 1)!$ for $\Gamma(x)$ is used, even for noninteger values of x , and the gamma function is also known as the factorial function.

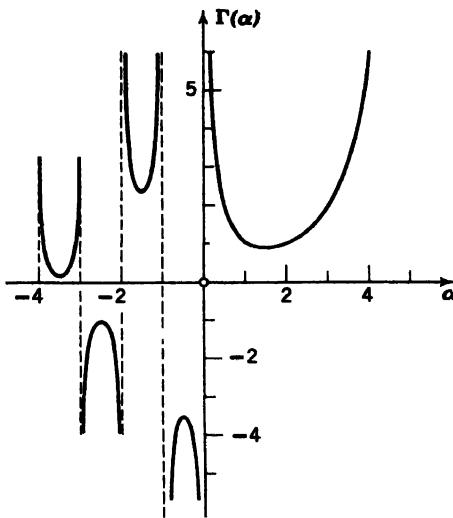


Fig. 373. Gamma function.

By repeated application of (3) we obtain

$$\Gamma(\alpha) = \frac{\Gamma(\alpha + 1)}{\alpha} = \frac{\Gamma(\alpha + 2)}{\alpha(\alpha + 1)} = \cdots = \frac{\Gamma(\alpha + k + 1)}{\alpha(\alpha + 1)(\alpha + 2) \cdots (\alpha + k)}$$

and we may use this relation

$$(6) \quad \Gamma(\alpha) = \frac{\Gamma(\alpha + k + 1)}{\alpha(\alpha + 1) \cdots (\alpha + k)} \quad (\alpha \neq 0, -1, -2, \dots)$$

for defining the gamma function for negative α ($\neq 0, -1, -2, \dots$), choosing for k the smallest integer such that $\alpha + k + 1 > 0$. Together with (1), this then gives a definition of $\Gamma(\alpha)$ for all α not equal to zero or a negative integer (Fig. 373).

Tables of $\Gamma(\alpha)$ for $1 \leq \alpha \leq 2$ have been computed (cf. Table 23 and Refs. [A7], [J2]), and for other values of α the values of $\Gamma(\alpha)$ can be obtained from these tables by using (3). A table $k!$ is included in Sec. 18.6.

It can be shown that the gamma function may also be represented as the limit of a product,³ namely, by the formula

$$(7) \quad \Gamma(\alpha) = \lim_{n \rightarrow \infty} \frac{n! n^\alpha}{\alpha(\alpha + 1)(\alpha + 2) \cdots (\alpha + n)} \quad (\alpha \neq 0, -1, \dots).$$

From (6) or (7) we see that, for complex α , the gamma function $\Gamma(\alpha)$ is a meromorphic function which has simple poles at $\alpha = 0, -1, -2, \dots$.

³ Cf. Ref. [A14]. Gauss used (7) as a definition of the gamma function; the integral (1) was introduced by Euler.

Since the gamma function increases very rapidly as α increases, it is sometimes convenient to use approximation formulas for $\Gamma(\alpha)$ when α is large. An important such formula is the Stirling⁴ formula

$$(8) \quad \Gamma(\alpha + 1) \approx \sqrt{2\pi\alpha} \left(\frac{\alpha}{e} \right)^{\alpha} \quad (\alpha \text{ a large positive number}),$$

where e is the base of the natural logarithm. The derivation of (8) (and similar more accurate but more complicated formulas) can be found in Ref. [A14] in Appendix 1.

Example 1. The following simple computation may illustrate the goodness of the approximation obtained by the Stirling formula. The values in the table (below) were computed by using (8) in the form

$$(9) \quad \ln k! \approx A_k + \frac{1}{2} \ln 2\pi \quad \text{where} \quad A_k = (k + \frac{1}{2}) \ln k - k$$

and a five-place table of natural logarithms. We see that with increasing k the relative error decreases while the absolute error increases. It can be shown that, as k approaches infinity the relative error approaches zero while the absolute error approaches infinity, a property which should be kept in mind when using the Stirling formula.

k	$\ln k$	A_k	Approximation		$k!$	Error	
			$\ln k!$	$k!$		abs.	rel.
4	1.386 29	2.238 31	3.157 25	23.5	24	0.5	2.0%
6	1.791 76	5.646 44	6.565 38	710	720	10	1.4%
8	2.079 44	9.675 24	10.594 18	39 900	40 320	$4 \cdot 10^2$	1.0%
10	2.302 59	14.177 20	15.096 14	3 600 000	3 628 800	$3 \cdot 10^4$	0.8%

The values of $\Gamma(k + \frac{1}{2})$ where k is an integer can easily be computed by using (3) and

$$(10) \quad \Gamma(\frac{1}{2}) = \sqrt{\pi}.$$

To prove (10), we start from (1) and set $t = u^2$. Then $dt = 2u \, du$ and

$$(10^*) \quad \Gamma(\frac{1}{2}) = \int_0^\infty e^{-t} t^{-\frac{1}{2}} dt = 2 \int_0^\infty e^{-u^2} du.$$

Squaring on both sides, writing v instead of u in the second integral and then writing the product of the two integrals as a double integral, we find

$$[\Gamma(\frac{1}{2})]^2 = 4 \int_0^\infty e^{-u^2} du \int_0^\infty e^{-v^2} dv = 4 \int_0^\infty \int_0^\infty e^{-(u^2+v^2)} \, du \, dv$$

We introduce polar coordinates r and ϕ by setting

$$u = r \cos \phi, \quad v = r \sin \phi.$$

⁴ JAMES STIRLING (1692–1770), English mathematician.

The element of area in polar coordinates is $r dr d\phi$, and the limits 0 and ∞ for u and v correspond to the limits 0 and ∞ for r and 0 and $\pi/2$ for ϕ . Thus,

$$[\Gamma(\frac{1}{2})]^2 = 4 \int_0^{\pi/2} \int_0^\infty e^{-r^2} r dr d\phi = 4 \cdot \frac{\pi}{2} \int_0^\infty e^{-r^2} r dr = 2\pi(-\frac{1}{2})e^{-\infty} = \pi.$$

By taking square roots on both sides we obtain (10).

We finally mention the so-called multiplication formula^b

$$(11) \quad \Gamma(\alpha)\Gamma(1-\alpha) = \frac{\pi}{\sin \pi\alpha} \quad (\alpha \neq 0, \pm 1, \pm 2, \dots)$$

We shall now consider functions which are closely related to the gamma function.

The functions

$$(12) \quad P(\alpha, x) = \int_0^x e^{-t} t^{\alpha-1} dt \quad \text{and} \quad Q(\alpha, x) = \int_x^\infty e^{-t} t^{\alpha-1} dt \quad (\alpha > 0)$$

are called the **incomplete gamma functions**. These functions of the two variables x and α occur occasionally in applications and play a certain role in mathematical statistics; we shall see later that they are related to other important special functions. From (1) we obtain

$$(13) \quad \Gamma(\alpha) = P(\alpha, x) + Q(\alpha, x).$$

Euler's **beta function** $B(x, y)$ is defined by the integral

$$(14) \quad B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt \quad (x > 0, y > 0).$$

It can be represented in terms of gamma functions as follows:

$$(15) \quad B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}. \checkmark$$

To derive (15), we start from

$$\begin{aligned} \Gamma(x)\Gamma(y) &= \int_0^\infty e^{-t} t^{x-1} dt \int_0^\infty e^{-u} u^{y-1} du \\ &= \int_0^\infty \int_0^\infty e^{-(t+u)} t^{x-1} u^{y-1} dt du. \end{aligned}$$

We introduce new variables of integration r and ϕ by setting

$$(16) \quad t = r \sin^2 \phi, \quad u = r \cos^2 \phi.$$

Expressing the integrand in terms of r and ϕ and replacing $dt du$ by $dr d\phi$ gives the absolute value of the Jacobian (cf. Sec. 6.3)

$$\begin{aligned} J(r, \phi) &= \begin{vmatrix} \frac{\partial t}{\partial r} & \frac{\partial t}{\partial \phi} \\ \frac{\partial u}{\partial r} & \frac{\partial u}{\partial \phi} \end{vmatrix} = \begin{vmatrix} \sin^2 \phi & 2r \sin \phi \cos \phi \\ \cos^2 \phi & -2r \cos \phi \sin \phi \end{vmatrix} = -2r \cos \phi \sin \phi \end{aligned}$$

For the proof see, e.g., Ref. [G5] in Appendix 1.

corresponding to (16), we obtain

$$(17) \quad \begin{aligned} \Gamma(x)\Gamma(y) &= 2 \int_0^{\pi/2} \int_0^\infty e^{-r} (r \sin^2 \phi)^{x-1} (r \cos^2 \phi)^{y-1} r \cos \phi \sin \phi dr d\phi \\ &= 2 \int_0^\infty e^{-t} t^{x+y-1} dt \int_0^{\pi/2} \sin^{2x-1} \phi \cos^{2y-1} \phi d\phi. \end{aligned}$$

The first integral on the right is $\Gamma(x + y)$. Furthermore, by setting $t = \sin^2 \phi$ in (14) we have $1 - t = \cos^2 \phi$, $dt = 2 \sin \phi \cos \phi d\phi$ and, therefore,

$$(18) \quad B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt = 2 \int_0^{\pi/2} \underline{\sin^{2x-1} \phi \cos^{2y-1} \phi d\phi}.$$

Hence (17) assumes the form

$$\Gamma(x)\Gamma(y) = \Gamma(x + y)B(x, y),$$

and (15) is proved.

Example 2. There are various integrals which can be expressed in terms of beta functions. For instance, from (18), (15), and (10) it follows that

$$(19) \quad \int_0^{\pi/2} \cos^\alpha \phi d\phi = \frac{1}{2} B\left(\frac{1}{2}, \frac{1}{2}\alpha + \frac{1}{2}\right) = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{1}{2}\alpha + \frac{1}{2})}{2\Gamma(\frac{1}{2}\alpha + 1)} = \frac{\sqrt{\pi}}{2} \frac{\Gamma(\frac{1}{2}\alpha + \frac{1}{2})}{\Gamma(\frac{1}{2}\alpha + 1)}.$$

Table 23. Gamma Function.

α	$\Gamma(\alpha)$	α	$\Gamma'(\alpha)$	α	$\Gamma'(\alpha)$	α	$\Gamma'(\alpha)$	α	$\Gamma(\alpha)$
1.00	1.000 000	1.20	0.918 169	1.40	0.887 264	1.60	0.893 515	1.80	0.931 384
1.02	0.988 844	1.22	0.913 106	1.42	0.886 356	1.62	0.895 924	1.82	0.936 845
1.04	0.978 438	1.24	0.908 521	1.44	0.885 805	1.64	0.898 642	1.84	0.942 612
1.06	0.968 744	1.26	0.904 397	1.46	0.885 604	1.66	0.901 668	1.86	0.948 687
1.08	0.959 725	1.28	0.900 718	1.48	0.885 747	1.68	0.905 001	1.88	0.955 071
1.10	0.951 351	1.30	0.897 471	1.50	0.886 227	1.70	0.908 639	1.90	0.961 766
1.12	0.943 590	1.32	0.894 640	1.52	0.887 039	1.72	0.912 581	1.92	0.968 774
1.14	0.936 416	1.34	0.892 216	1.54	0.888 178	1.74	0.916 826	1.94	0.976 099
1.16	0.929 803	1.36	0.890 185	1.56	0.889 639	1.76	0.921 375	1.96	0.983 743
1.18	0.923 728	1.38	0.888 537	1.58	0.891 420	1.78	0.926 227	1.98	0.991 708
1.20	0.918 169	1.40	0.887 264	1.60	0.893 515	1.80	0.931 384	2.00	1.000 000

Problems

Using Table 23, compute

1. $\Gamma(2.3)$ 2. $\Gamma(4.5)$ 3. $\Gamma(-0.4)$ 4. $\Gamma(-1.6)$ 5. $\Gamma(-2.7)$

6. Express $\Gamma(k + \frac{1}{2})$ where k is a positive integer, in terms of $\Gamma(\frac{1}{2})$ and compute $\Gamma(k + \frac{1}{2})$ for $k = 1, 2, 3, 4$.

Compute

7. $B(1, 2)$ 8. $B(3, 3)$ 9. $B(2, 4)$ 10. $B(2.5, 1.5)$ 11. $B(1.9, 0.6)$

12. Prove that the last integral in (2) exists for all real α and all $c > 0$.

13. Derive (3) from (7).

14. Prove that the residue of $\Gamma(\alpha)$ at $\alpha = -k$ (k a non-negative integer) is $(-1)^k/k!$.

Show that

15. $\int_0^{\pi/2} \cos^{2n} \phi d\phi = \frac{\pi}{2} \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots (2n)} \quad (n = 1, 2, \dots)$

16. $\int_0^{\pi/2} \cos^{2n+1} \phi d\phi = \frac{2^{2n}(n!)^2}{(2n+1)!} \quad (n = 0, 1, \dots)$

17. $\int_0^{\pi/2} \sin^2 \phi \cos^2 \phi d\phi = \frac{\pi}{16}$ ✓

18. $\int_0^{\pi/2} \sin^5 \phi \cos^7 \phi d\phi = \frac{1}{120}$

19. $\int_{-1}^1 (1-t^2)^n dt = \frac{2^{n+1} n!}{1 \cdot 3 \cdot 5 \cdots (2n+1)} \quad (n = 0, 1, \dots)$ [Set $t = \sin \phi$ in (19).]

20. Express $\int_0^{\pi/2} \sin^\alpha \phi d\phi$ ($\alpha > 0$) in terms of gamma functions. Discuss in particular the cases $\alpha = 2n$ and $\alpha = 2n + 1$, where $n = 0, 1, \dots$.

17.2 Error Function. Fresnel Integrals. Sine and Cosine Integrals

The function e^{-x^2} is an even function whose graph, the so-called *bell-shaped curve*, is shown in Fig. 374. This function and its integral frequently appear in applied mathematics. The function

$$(1) \quad \text{erf } x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

is called the **error function** (Fig. 375).

From (10) and (10*) in the last section we see that

$$(2) \quad \int_0^\infty e^{-t^2} dt = \frac{1}{2}\Gamma(\frac{1}{2}) = \frac{1}{2}\sqrt{\pi}. \quad \checkmark$$

From this and (1) it follows that

$$(3) \quad \text{erf } \infty = \lim_{x \rightarrow \infty} \text{erf } x = 1,$$

and this is the reason for choosing the “normalizing factor” $2/\sqrt{\pi}$ in (1).

The error function cannot be represented in terms of finitely many elementary functions. Tables of function values of $\text{erf } x$ are available (see Table 24);

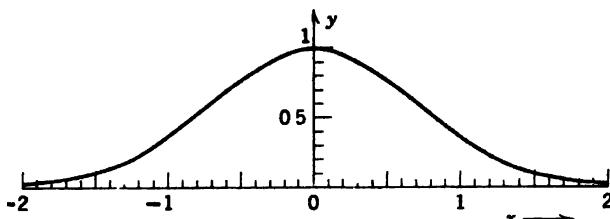
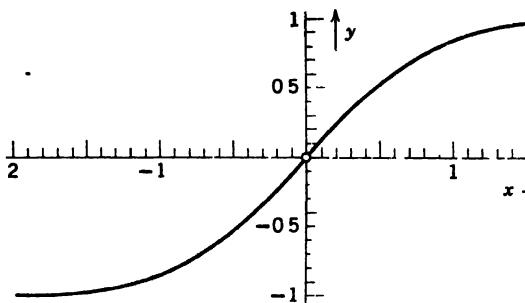
Fig. 374. Graph of $y = e^{-x^2}$ ("bell-shaped curve").

Fig. 375. Error function.

cf. also Ref. [J8] in Appendix 1). For computing $\operatorname{erf} x$ for small values of $|x|$ we may use the Maclaurin series

$$(4) \quad \operatorname{erf} x = \frac{2}{\sqrt{\pi}} \left(x - \frac{x^3}{1! 3} + \frac{x^5}{2! 5} - \frac{x^7}{3! 7} + \dots \right), \checkmark$$

which is obtained by integrating the Maclaurin series of e^{-x^2} term by term. For large values of x we may use the asymptotic expansion of $\operatorname{erf} x$, which will be derived in the next section.

The function

$$(5) \quad \operatorname{erfc} x = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt \quad \checkmark$$

is called the complementary error function. From (3) it follows that

$$(6) \quad \operatorname{erfc} x = 1 - \operatorname{erf} x.$$

We finally mention some other important integrals which cannot be evaluated in terms of finitely many elementary functions.

The Fresnel integrals⁶ (Fig. 376)

$$(7) \quad C(x) = \int_0^x \cos(t^2) dt, \quad S(x) = \int_0^x \sin(t^2) dt$$

⁶ AUGUSTIN FRESNEL (1788–1827), French physicist and mathematician. For tables see Ref. [A6], [A7]; cf. also *Acta mathematica*, vol. 85 (1951), p. 180, and vol. 89 (1953), p. 130.

Table 24. Error Function and Its Derivative.

x	$\operatorname{erf} x$	$(2/\sqrt{\pi})e^{-x^2}$	x	$\operatorname{erf} x$	$(2/\sqrt{\pi})e^{-x^2}$
0.00	0.000 000	1.128 379	1.00	0.842 701	0.415 107
0.05	0.056 372	1.125 562	1.1	0.880 205	0.336 480
0.10	0.112 463	1.117 152	1.2	0.910 314	0.267 344
0.15	0.167 996	1.103 274	1.3	0.934 008	0.208 208
0.20	0.222 703	1.084 135	1.4	0.952 285	0.158 942
0.25	0.276 326	1.060 014	1.5	0.966 105	0.118 930
0.30	0.328 627	1.031 261	1.6	0.976 348	0.087 229
0.35	0.379 382	0.998 284	1.7	0.983 790	0.062 711
0.40	0.428 392	0.961 541	1.8	0.989 091	0.044 192
0.45	0.475 482	0.921 532	1.9	0.992 790	0.030 525
0.50	0.520 500	0.878 783	2.0	0.995 322	0.020 667
0.55	0.563 323	0.833 837	2.2	0.998 137	0.008 922
0.60	0.603 856	0.787 243	2.4	0.999 311	0.003 556
0.65	0.642 029	0.739 547	2.6	0.999 764	0.001 308
0.70	0.677 801	0.691 275	2.8	0.999 925	0.000 444
0.75	0.711 156	0.642 931	3.0	0.999 978	0.000 139
0.80	0.742 101	0.594 986	3.2	0.999 994	0.000 040
0.85	0.770 668	0.547 870	3.4	0.999 998	0.000 011
0.90	0.796 908	0.501 969	3.6	1.000 000	0.000 003
0.95	0.820 891	0.457 619	3.8	1.000 000	0.000 001
1.00	0.842 701	0.415 107	4.0	1.000 000	0.000 000

Table 25. Sine and Cosine Integrals.

x	$\text{Si}(x)$	$\text{ci}(x)$	x	$\text{Si}(x)$	$\text{ci}(x)$	x	$\text{Si}(x)$	$\text{ci}(x)$
0.0	0.0000	∞	2.0	1.6054	-0.4230	5	1.5499	0.1900
0.2	0.1996	1.0422	2.2	1.6876	-0.3751	6	1.4247	0.0681
0.4	0.3965	0.3788	2.4	1.7525	-0.3173	7	1.4546	-0.0767
0.6	0.5881	0.0223	2.6	1.8004	-0.2533	8	1.5742	-0.1224
0.8	0.7721	-0.1983	2.8	1.8321	-0.1865	9	1.6650	-0.0554
1.0	0.9461	-0.3374	3.0	1.8487	-0.1196	10	1.6583	0.0455
1.2	1.1080	-0.4205	3.2	1.8514	-0.0553	11	1.5783	0.0896
1.4	1.2562	-0.4620	3.4	1.8419	0.0045	12	1.5050	0.0498
1.6	1.3892	-0.4717	3.6	1.8219	0.0580	13	1.4994	-0.0268
1.8	1.5058	-0.4568	3.8	1.7934	0.1038	14	1.5562	-0.0694
2.0	1.6054	-0.4230	4.0	1.7582	0.1410	15	1.6182	-0.0463

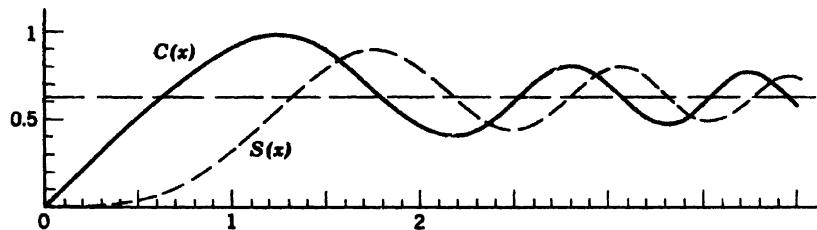


Fig. 376. Fresnel integrals.

and their complementary functions

$$(8) \quad c(x) = \int_x^{\infty} \cos(t^2) dt, \quad s(x) = \int_x^{\infty} \sin(t^2) dt$$

appear in various applications, for example, in optics and in antenna theory. We have shown that

$$(9) \quad C(\infty) = \int_0^{\infty} \cos(t^2) dt = \frac{1}{2}\sqrt{\frac{\pi}{2}}, \quad S(\infty) = \int_0^{\infty} \sin(t^2) dt = \frac{1}{2}\sqrt{\frac{\pi}{2}},$$

see (17) in Sec. 15.4. Therefore,

$$(10) \quad c(x) = \frac{1}{2}\sqrt{\frac{\pi}{2}} - C(x), \quad s(x) = \frac{1}{2}\sqrt{\frac{\pi}{2}} - S(x).$$

The sine integral (Fig. 377) is defined as follows:

$$(11) \quad Si(x) = \int_0^x \frac{\sin t}{t} dt.$$

Its complementary function is

$$(12) \quad si(x) = \int_x^{\infty} \frac{\sin t}{t} dt.$$

Since

$$(13) \quad Si(\infty) = \int_0^{\infty} \frac{\sin t}{t} dt = \frac{\pi}{2}$$

(see Ex. 4, Sec. 15.4), we have

$$(14) \quad si(x) = \frac{\pi}{2} - Si(x).$$

The function

$$(15) \quad ci(x) = \int_x^{\infty} \frac{\cos t}{t} dt$$

is called the cosine integral.

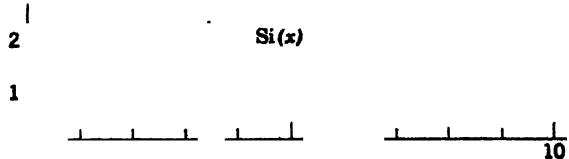


Fig. 377. Sine integral.

Problems

1. Show that $\operatorname{erf} x$, $C(x)$, $S(x)$, and $\operatorname{Si}(x)$ are odd functions.
2. Derive (4). Why is term-by-term integration permissible?

Find the Maclaurin series of

3. $C(x)$

4. $S(x)$

5. $\operatorname{Si}(x)$

6. Graph $\operatorname{erfc} x$ and $\operatorname{si}(x)$. Carry out the derivation of (6).

7. Determine the location of the maxima and minima of $C(x)$, $S(x)$, and $\operatorname{Si}(x)$.

Show that

8. $C(x) = \frac{1}{2} \int_0^{x^2} \frac{\cos \tau}{\sqrt{\tau}} d\tau$

9. $S(x) = \frac{1}{2} \int_0^{x^2} \frac{\sin \tau}{\sqrt{\tau}} d\tau$

10. $\operatorname{erf} x = P(\frac{1}{2}, x^2)/\sqrt{\pi}$, $\operatorname{erfc} x = Q(\frac{1}{2}, x^2)/\sqrt{\pi}$ (cf. Sec. 17.1)

11. $C(x) = \frac{\sqrt{\pi}}{4} [e^{i\pi/4} \operatorname{erf}(xe^{-i\pi/4}) + e^{-i\pi/4} \operatorname{erf}(xe^{i\pi/4})]$

12. $S(x) = \frac{\sqrt{\pi}}{4i} [e^{i\pi/4} \operatorname{erf}(xe^{-i\pi/4}) - e^{-i\pi/4} \operatorname{erf}(xe^{i\pi/4})]$

13. Show that $\operatorname{Si}(z)$ does not have real zeros (except at $z = 0$).

14. The integral in (11) with $\sin t$ replaced by $\cos t$ does not exist. Why?

15. Show that $C(z)$, $S(z)$, and $\operatorname{Si}(z)$ are entire functions of $z = x + iy$.

17.3 Asymptotic Expansions

Asymptotic expansions are (in general divergent) series which are of great practical importance for computing values of a function $f(x)$ for large x . It is clear that the Maclaurin series of $f(x)$, if it exists and converges for large x , is not suitable for that purpose, because the number of terms to be used for obtaining a prescribed accuracy increases rapidly as x increases. For a Taylor series with center at a , where a is large, the situation will be similar for all those x which are far apart from a . We shall see that the larger x is, the fewer terms of an asymptotic expansion we need for obtaining a required accuracy. On the other hand, the accuracy is limited, and it decreases as x decreases so that asymptotic expansions can be used for large x only.

The variables and functions to be considered in this section are assumed to be real.

A series of the form

$$c_0 + \frac{c_1}{x} + \frac{c_2}{x^2} + \dots \quad (c_0, c_1, \dots \text{ constant})$$

(which need not converge for any value of x) is called an **asymptotic expansion, or asymptotic series, of a function $f(x)$ which is defined for every sufficiently**

large value of x if, for every fixed $n = 0, 1, 2, \dots$,

$$(1) \quad \left[f(x) - \left(c_0 + \frac{c_1}{x} + \frac{c_2}{x^2} + \cdots + \frac{c_n}{x^n} \right) \right] x^n \rightarrow 0 \quad \text{as } x \rightarrow \infty,$$

and we shall then write

$$f(x) \sim c_0 + \frac{c_1}{x} + \frac{c_2}{x^2} + \cdots.$$

If a function $f(x)$ has an asymptotic expansion, then this expansion is unique, because its coefficients c_0, c_1, \dots are uniquely determined by (1). In fact, from (1) we obtain

$$(1^*) \quad \begin{aligned} f(x) - c_0 &\rightarrow 0 & \text{or} & \quad c_0 = \lim_{x \rightarrow \infty} f(x), \\ \left[f(x) - c_0 - \frac{c_1}{x} \right] x &\rightarrow 0 & \text{or} & \quad c_1 = \lim_{x \rightarrow \infty} [f(x) - c_0]x, \quad \text{etc.} \end{aligned}$$

On the other hand, different functions may have the same asymptotic expansion. In fact, let $f(x) = e^{-x}$. Then, since $e^{-x} \rightarrow 0$, $xe^{-x} \rightarrow 0$, etc., we see from (1^{*}) that $c_0 = 0$, $c_1 = 0$, etc. Hence,

$$e^{-x} \sim 0 + \frac{0}{x} + \cdots.$$

Thus, if $g(x)$ has an asymptotic expansion, then $g(x) + e^{-x}$ has certainly the same asymptotic expansion.

For applications it is advantageous to extend the definition by writing

$$f(x) \sim g(x) + h(x) \left(c_0 + \frac{c_1}{x} + \frac{c_2}{x^2} + \cdots \right)$$

whenever

$$\frac{f(x) - g(x)}{h(x)} \sim c_0 + \frac{c_1}{x} + \frac{c_2}{x^2} + \cdots$$

in the sense of the preceding definition.

Only in rare cases may we determine the coefficients of an asymptotic expansion directly from (1^{*}). In general, other methods will be more suitable, for example, successive integration by parts. We shall illustrate this procedure by Ex. 1, and we shall also compare the typical behavior of the terms of an asymptotic series with that of the terms of a Maclaurin series.

Example 1 (Asymptotic series of the error function). We have

$$\operatorname{erf} x = 1 - \operatorname{erfc} x,$$

and by setting $t^2 = \tau$, $dt = d\tau/2\sqrt{\tau}$,

$$\operatorname{erfc} x = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt = \frac{1}{\sqrt{\pi}} \int_x^\infty e^{-\tau} \tau^{-1/2} d\tau.$$

Repeated integration by parts will lead to integrals of the form

$$(2) \quad F_n(x) = \int_{x^2}^{\infty} e^{-\tau} \tau^{-(2n+1)/2} d\tau, \quad n = 0, 1, \dots$$

Note that $\text{erfc } x = F_0(x)/\sqrt{\pi}$. By integration by parts we obtain

$$F_n(x) = -e^{-x} x^{-(2n+1)/2} \Big|_{x^2}^{\infty} - \frac{2n+1}{2} \int_{x^2}^{\infty} e^{-\tau} \tau^{-(2n+3)/2} d\tau.$$

The integral on the right is $F_{n+1}(x)$ and, therefore,

$$F_n(x) = \frac{1}{x^{2n+1}} e^{-x^2} - \frac{2n+1}{2} F_{n+1}(x), \quad n = 0, 1, \dots$$

If we multiply by e^{x^2} on both sides, this recurrence relation becomes

$$e^{x^2} F_n(x) = \frac{1}{x^{2n+1}} - \frac{2n+1}{2} e^{x^2} F_{n+1}(x).$$

Repeated application of this formula yields

$$\begin{aligned} (3) \quad e^{x^2} F_0(x) &= \frac{1}{x} - \frac{1}{2} e^{x^2} F_1(x) \\ &= \frac{1}{x} - \frac{1}{2x^3} + \frac{1}{2} \cdot \frac{3}{2} e^{x^2} F_2(x) \\ &\quad \dots \dots \dots \\ &= \left[\frac{1}{x} - \frac{1}{2x^3} + \frac{1 \cdot 3}{2^2 x^5} - + \dots + (-1)^{n-1} \frac{1 \cdot 3 \cdots (2n-3)}{2^{n-1} x^{2n-1}} \right] \\ &\quad + (-1)^n \frac{1 \cdot 3 \cdots (2n-1)}{2^n} e^{x^2} F_n(x). \end{aligned}$$

We show that the series obtained in this way is an asymptotic expansion,

$$(4) \quad e^{x^2} F_0(x) \sim \frac{1}{x} - \frac{1}{2x^3} + \frac{1 \cdot 3}{2^2 x^5} - + \dots .$$

Let S_{2n-1} denote the expression in brackets in (3). Then from (3) we obtain

$$(5) \quad [e^{x^2} F_0(x) - S_{2n-1}] x^{2n-1} = K_n e^{x^2} x^{2n-1} F_n(x)$$

where

$$K_n = (-1)^n \frac{1 \cdot 3 \cdots (2n-1)}{2^n}.$$

We have to show that for each fixed $n = 1, 2, \dots$ the expression on the right approaches zero as x approaches infinity. In (2),

$$\frac{1}{\tau^{(2n+1)/2}} \leq \frac{1}{x^{2n+1}} \quad \text{for all } \tau \geq x^2.$$

Hence we obtain the inequality

$$(6) \quad F_n(x) = \int_{x^2}^{\infty} \frac{e^{-\tau}}{\tau^{(2n+1)/2}} d\tau < \frac{1}{x^{2n+1}} \int_{x^2}^{\infty} e^{-\tau} d\tau = \frac{e^{-x^2}}{x^{2n+1}}.$$

From this we immediately see that

$$|K_n| e^{x^2} x^{3n-1} F_n(x) < \frac{|K_n|}{x} \rightarrow 0 \quad (x \rightarrow \infty).$$

This proves that the series in (4) is an asymptotic expansion of the function on the left-hand side of (4). Since

$$\operatorname{erf} x = 1 - \operatorname{erfc} x = 1 - \frac{F_0(x)}{\sqrt{\pi}}$$

it follows that the desired asymptotic series of the error function is

$$(7) \quad \operatorname{erf} x \sim 1 - \frac{1}{\sqrt{\pi}} e^{-x^2} \left(\frac{1}{x} - \frac{1}{2x^3} + \frac{1 \cdot 3}{2^2 x^5} - \frac{1 \cdot 3 \cdot 5}{2^3 x^7} + \dots \right).$$

For large x we thus have the simple approximation

$$(7*) \quad \operatorname{erf} x \approx 1 - \frac{1}{\sqrt{\pi} x} e^{-x^2}.$$

We now want to illustrate the use of (7) in numerical computation. From (5) and (6) it follows that

$$|e^{x^2} F_0(x) - S_{2n-1}| = \frac{1 \cdot 3 \cdots (2n-1)}{2^n} e^{x^2} F_n(x) < \frac{1 \cdot 3 \cdots (2n-1)}{2^n} \frac{1}{x^{2n+1}},$$

and for sufficiently large x the expression on the right is very small. This shows that for such x , S_{2n-1} is a very good approximation to $e^{x^2} F_0(x)$, and so, for large x , the error function can be computed very accurately by means of (7). It turns out that even for relatively small $|x|$ the results are surprisingly accurate.

For instance, let us take $x = 2$ and show that by taking the sum of a suitable number of terms in (7) we obtain three digit accuracy of $\operatorname{erf} 2 (= 0.99532)$. Table 26 shows that the terms in (7) first decrease in absolute value and then (for $n > 5$) again increase (cf. Fig. 378). This is typical for an asymptotic expansion. In general, the greatest accuracy is obtained by taking the sum of the terms up to the smallest in absolute value. In our case, this is the sum of the first five terms (corresponding to $n = 1, \dots, 5$); by taking more terms the error again increases and, therefore, the accuracy is limited to three digits. This limited accuracy is in contrast to the situation in the case of a Maclaurin series where, for every fixed x for which the series converges, we can reach any degree of accuracy by taking the sum of sufficiently many terms. Hence, for fixed x , when using an asymptotic expansion we have a limited degree of accuracy. However, the advantage of the asymptotic expansion is obvious.

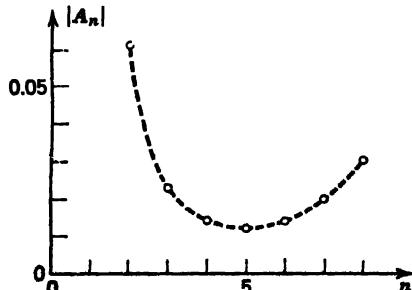


Fig. 378. Absolute values of the terms A_n of the asymptotic expansion (7) for $x = 2$.

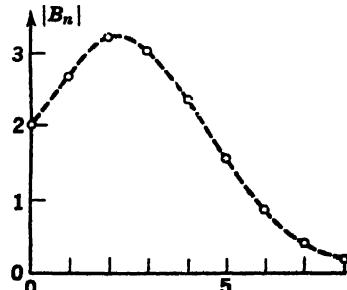


Fig. 379. Absolute values of the terms of the Maclaurin series (8) for $x = 2$.

In our case we need only two terms (the constant term and the term involving $1/x$) to obtain the first three decimals of $\text{erf } 2$, and the error of this approximation is $1/2$ unit of the third decimal (0.000 49). Figure 379 shows the absolute values of the first few terms of the Maclaurin series

$$(8) \quad \frac{\sqrt{\pi}}{2} \text{erf } x = x - \frac{x^3}{1!3} + \frac{x^5}{2!5} - \frac{x^7}{3!7} + \dots$$

when $x = 2$. The values of these terms $B_n = (-1)^n 2^{2n+1}/n! (2n+1)$ for $n = 0, \dots, 13$ are

$$\begin{array}{cccccccc} 2.000\ 00 & -2.666\ 67 & 3.200\ 00 & -3.047\ 62 & 2.370\ 37 & -1.551\ 52 & 0.875\ 21 \\ -0.433\ 44 & 0.191\ 22 & -0.076\ 04 & 0.027\ 52 & -0.009\ 14 & 0.002\ 80 & -0.000\ 80 \end{array}$$

We see that to compute the first three decimals of $\text{erf } 2$ from (8) we must use the sum of the first fourteen terms, the last one of which contains r^{27} . If four or more correct digits of $\text{erf } 2$ are needed, then we cannot use the asymptotic expansion, as we see from the error in Table 26.

Table 26. Computation of $\text{erf } x$ for $x = 2$ by Means of the Asymptotic Expansion.

n	$A_1 = 1/x$ $A_n = (-1)^{n-1} \frac{1 \cdot 3 \cdots (2n-3)}{2^n \cdot 1 \cdot 3 \cdots (2n-1)}$	$1 - \frac{1}{\sqrt{\pi}} e^{-x^2} \sum_{j=1}^n A_j$	Absolute value of error	Error bound $k(2, n)$
1	0.500 000	0.994 83	0.000 49	0.000 65
2	-0.062 500	0.995 48	0.000 16	0.000 25
3	0.023 438	0.995 24	0.000 08	0.000 16
4	-0.014 648	0.995 39	0.000 07	0.000 14
5	0.012 817	0.995 26	0.000 06	0.000 15
6	-0.014 420	0.995 40	0.000 08	0.000 21
7	0.019 827	0.995 20	0.000 12	0.000 34
8	-0.032 219	0.995 53	0.000 21	0.000 63

We mention that we can obtain a bound for the absolute value of the error of our approximate values in Table 26 as follows. From (3) and

$$\text{erf } x = 1 - \text{erfc } x = 1 - \frac{1}{\sqrt{\pi}} F_0(x)$$

we see that the absolute value of the error is

$$\frac{1 \cdot 3 \cdots (2n-1)}{2^n \sqrt{\pi}} |F_n(x)|.$$

From (6) it follows that this expression is less than

$$k(x, n) = \frac{1 \cdot 3 \cdots (2n-1)}{2^n \sqrt{\pi}} \frac{e^{-x^2}}{x^{2n+1}} \quad (x > 0, n = 1, 2, \dots),$$

and this is the desired bound. Numerical values for $x = 2$ are shown in Table 26.

To understand completely our result concerning the asymptotic expansion, the reader should keep in mind that $x = 2$ is relatively small; for $x = 5$, for example, we obtain from (7*)

$$\operatorname{erf} 5 \approx 1 - \frac{1}{\sqrt{x}} \frac{e^{-25}}{5} = 0.99999\ 99999\ 98433;$$

the first thirteen decimals are correct, the error is three units of the fourteenth decimal. However, we wish to warn the reader that the present case may be misleading; for other asymptotic expansions a similarly favourable situation may arise not at $x = 2$ or 5, but only for much larger x , perhaps for $x = 20$ or $x = 100$.

Problems

1. Draw figures similar to Figs. 378 and 379 for $\operatorname{erf} x$ when $x = 1$.
2. Same task as in Prob. 1 when $x = 3$.
3. Find a method for determining the smallest term in absolute value of the asymptotic series $c_0 + c_1/x + \dots$.
4. Why is it important to determine the smallest term in absolute value of an asymptotic series to be used for computations?
5. Which term of (7) is the smallest in absolute value when $x = 1$? When $x = 2$? When $x = 3$?
6. Find a method for determining the largest term in absolute value of a Maclaurin series $a_0 + a_1x + a_2x^2 + \dots$, where x is fixed.
7. Which term of the Maclaurin series of e^x is the largest when $x = 2$? When $x = 3$? When $x = n$?
8. Compute a table of $\operatorname{erf} x$ for $x = 0, 0.5, 1.0, \dots, 5.0$ (two decimal places).
9. Show that the series $1 + \frac{1}{x} + \frac{1}{2!x^2} + \dots$ [which converges for $|x| > 0$ and represents $e^{1/x}$] is an asymptotic expansion of $e^{1/x}$.
10. Show that $\sin \frac{1}{x} \sim \frac{1}{x} - \frac{1}{3!x^3} + \frac{1}{5!x^5} - + \dots$

17.4 Further Properties of Asymptotic Expansions

For practical applications it is useful to know that asymptotic series may be added, multiplied, and under certain restrictions also integrated and differentiated term by term. Let us formulate these properties in a precise manner.

Theorem 1. If

$$f(x) \sim a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \dots \quad \text{and} \quad g(x) \sim b_0 + \frac{b_1}{x} + \frac{b_2}{x^2} +$$

then the function $Af + Bg$, where A and B are constants, has the asymptotic expansion

$$(1) \quad Af(x) + Bg(x) \sim Aa_0 + Bb_0 + \frac{Aa_1 + Bb_1}{x} + \frac{Aa_2 + Bb_2}{x^2} + \dots$$

and the function fg has the asymptotic expansion

$$(2) \quad f(x)g(x) \sim c_0 + \frac{c_1}{x} + \frac{c_2}{x^2} + \dots$$

whose coefficients are given by the formula

$$c_n = a_0 b_n + a_1 b_{n-1} + \dots + a_n b_0.$$

Proof. The simple proof of (1) is left to the reader. We prove the last statement. We must show that for any fixed non-negative integer n ,

$$(3) \quad (fg - S_n)x^n \rightarrow 0 \quad \text{as } x \rightarrow \infty,$$

where

$$S_n(x) = c_0 + \frac{c_1}{x} + \dots + \frac{c_n}{x^n}$$

with the coefficients c_0, \dots, c_n given in the theorem.

We choose an arbitrary fixed n and write

$$f(x) = s_n(x) + \frac{h(x)}{x^n} \quad \text{where} \quad s_n(x) = a_0 + \frac{a_1}{x} + \dots + \frac{a_n}{x^n}.$$

Then we immediately have

$$[f(x) - s_n(x)]x^n = h(x).$$

From this and the definition of an asymptotic expansion it follows that $h(x)$ must approach zero as $x \rightarrow \infty$. Similarly, if we write

$$g(x) = s_n^*(x) + \frac{l(x)}{x^n} \quad \text{where} \quad s_n^*(x) = b_0 + \frac{b_1}{x} + \dots + \frac{b_n}{x^n},$$

then $l(x) \rightarrow 0$ as $x \rightarrow \infty$. From these representations we obtain

$$fg = s_n s_n^* + \frac{h + l}{x^n} + \frac{hl}{x^{2n}}.$$

By multiplying term by term and collecting like powers of x we can readily verify that

$$s_n s_n^* = S_n + T_n$$

where T_n is the sum of terms involving the powers $1/x^{n+1}, \dots, 1/x^{2n}$. Clearly, $x^n T_n \rightarrow 0$ as $x \rightarrow \infty$. Consequently, in (3),

$$fg - S_n = T_n + \frac{h + l}{x^n} + \frac{hl}{x^{2n}}.$$

We multiply by x^n on both sides. Then the resulting expression on the right approaches zero as $x \rightarrow \infty$, because $x^n T_n \rightarrow 0$, $l \rightarrow 0$, and $h \rightarrow 0$. Hence, the expression on the left-hand side of (3) approaches zero, and the proof is complete.

We now consider the term-by-term integration of asymptotic expansions.

Theorem 2. Let $f(x)$ be continuous for all sufficiently large x , and let

$$f(x) \sim \frac{c_2}{x^2} + \frac{c_3}{x^3} + \dots$$

Then for those x ,

$$(4) \quad \int_x^\infty f(t) dt \sim \frac{c_2}{x} + \frac{c_3}{2x^2} + \frac{c_4}{3x^3} + \dots$$

Proof. Let $F(x)$ denote the integral in (4) and let $S_{n-1}(x)$ denote the integral of

$$s_n(x) = \frac{c_2}{x^2} + \dots + \frac{c_n}{x^n},$$

that is,

$$S_{n-1}(x) = \int_x^\infty s_n(t) dt = \frac{c_2}{x} + \frac{c_3}{2x^2} + \dots + \frac{c_n}{(n-1)x^{n-1}}.$$

By definition of an asymptotic expansion, for every $n = 0, 1, \dots$,

$$|f(x) - s_n(x)| x^n \rightarrow 0 \quad \text{as } x \rightarrow \infty.$$

Because of continuity of f this implies that, given any $\epsilon > 0$ we can find an x_0 such that for all $x > x_0$,

$$|f(x) - s_n(x)| x^n < \epsilon \quad \text{or} \quad |f(x) - s_n(x)| < \frac{\epsilon}{x^n}.$$

From this we obtain for all $x > x_0$

$$\begin{aligned} |F(x) - S_{n-1}(x)| &= \left| \int_x^\infty f(t) dt - \int_x^\infty s_n(t) dt \right| = \left| \int_x^\infty [f(t) - s_n(t)] dt \right| \\ &\leq \int_x^\infty |f(t) - s_n(t)| dt < \epsilon \int_x^\infty \frac{dt}{t^n} = \frac{\epsilon}{(n-1)x^{n-1}}. \end{aligned}$$

By multiplying on both sides by the positive quantity x^{n-1} we have

$$|F(x) - S_{n-1}(x)| x^{n-1} < \frac{\epsilon}{n-1} \quad \text{when } x > x_0(\epsilon).$$

Since $\epsilon (>0)$ may be chosen as small as we please, it follows that

$$|F(x) - S_{n-1}(x)| x^{n-1} \rightarrow 0 \quad \text{as } x \rightarrow \infty,$$

and the proof is complete.

If $f(x)$ is continuous for all sufficiently large x and

$$f(x) \sim c_0 + \frac{c_1}{x} + \frac{c_2}{x^2} + \dots,$$

then we conclude from Theorem 2 that

$$(4^*) \quad \int_x^\infty \left[f(t) - c_0 - \frac{c_1}{t} \right] dt \sim \frac{c_2}{x} + \frac{c_3}{2x^2} + \frac{c_4}{3x^3} + \dots.$$

If $f(x)$ has an asymptotic expansion, it does not follow that the derivative $f'(x)$ has an asymptotic expansion. For example, from (1*) in the last section we obtain

$$f(x) = e^{-x} \sin(e^x) \sim 0 + \frac{0}{x} + \frac{0}{x^2} + \dots$$

but the derivative of $f(x)$, which is given by the formula

$$f'(x) = -e^{-x} \sin(e^x) + e^{-x} \cos(e^x) e^x = -f(x) + \cos(e^x),$$

does not have an asymptotic expansion. (Why?) However, if the derivative $f'(x)$ of a function $f(x)$ has an asymptotic expansion, it can be obtained by termwise differentiation of that of $f(x)$. In fact, the following theorem holds true.

Theorem 3. *If*

$$(5) \quad f(x) \sim c_0 + \frac{c_1}{x} + \frac{c_2}{x^2} + \dots$$

and $f(x)$ has a continuous derivative $f'(x)$ which has an asymptotic expansion, then this expansion is

$$(6) \quad f'(x) \sim -\frac{c_1}{x^2} - \frac{2c_2}{x^3} - \frac{3c_3}{x^4} - \dots.$$

Proof. By assumption,

$$(7) \quad f'(x) \sim a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \dots$$

and we have to show that the coefficients a_n are such that (6) and (7) are identical. We first show that $a_0 = 0$ and $a_1 = 0$. From (5) and the definition of an asymptotic expansion we have

$$(8) \quad (a) \quad \lim_{x \rightarrow \infty} f(x) = c_0, \quad (b) \quad \lim_{x \rightarrow \infty} [f(x) - c_0]x = c_1.$$

The corresponding relations for the asymptotic series (7) are

$$(9) \quad (a) \quad \lim_{x \rightarrow \infty} f'(x) = a_0, \quad (b) \quad \lim_{x \rightarrow \infty} [f'(x) - a_0]x = a_1.$$

We may relate f and f' by the formula

$$(10) \quad f(x) = \int_{x_0}^x f'(t) dt + k \quad [x_0 (> 0) \text{ and } k \text{ constant}].$$

From this and (8a) it follows that

$$\lim_{x \rightarrow \infty} \int_{x_0}^x f'(t) dt + k = c_0$$

Now (9a) shows that if a_0 were not zero, the limit of the integral would not exist. Hence $a_0 = 0$. Then (9b) becomes

$$(9b^*) \quad \lim_{x \rightarrow \infty} xf'(x) = a_1.$$

By definition this means that for any preassigned $\epsilon > 0$ and all sufficiently large x ,

$$(11) \quad a_1 - \epsilon < xf'(x) < a_1 + \epsilon \quad \text{or} \quad \frac{a_1 - \epsilon}{x} < f'(x) < \frac{a_1 + \epsilon}{x}.$$

From (8b) and (10) we obtain

$$\lim_{x \rightarrow \infty} \left(\int_{x_0}^x f'(t) dt + k - c_0 \right) x = c_1.$$

From (11) we see that for $a_1 \neq 0$ this limit does not exist. Hence $a_1 = 0$, and (7) becomes

$$f'(x) \sim \frac{a_2}{x^2} + \frac{a_3}{x^3} + \dots$$

Consequently, from (10) and Theorem 2 we now obtain

$$(12) \quad \begin{aligned} f(x) &= \int_{x_0}^{\infty} f'(t) dt - \int_x^{\infty} f'(t) dt + k \\ &\sim \int_{x_0}^{\infty} f'(t) dt + k - \frac{a_2}{x} - \frac{a_3}{x^2} - \dots \end{aligned}$$

the first integral on the right being a constant. If a given function has an asymptotic expansion, it is unique. We may thus compare corresponding terms in (5) and (12), finding $a_2 = -c_1$, $a_3 = -2c_2$, etc. With these coefficients the series (6) and (7) become identical, and the theorem is proved.

If we know that a function $f(x)$ satisfies the assumptions of Theorem 3 and is a solution of a first-order differential equation, then we may determine the coefficients of its asymptotic expansion by substituting (5) and (6) in the differential equation.

Example 1 (Exponential integral). Let

$$y = f(x) = e^x \int_x^{\infty} \frac{e^{-t}}{t} dt \quad (x > 0).$$

Then

$$y' = f'(x) = e^x \int_x^{\infty} \frac{e^{-t}}{t} dt - e^x \frac{e^{-x}}{x} = y - \frac{1}{x}.$$

Hence $f(x)$ satisfies the linear differential equation

$$(13) \quad y' - y + \frac{1}{x} = 0.$$

It may be proved directly that this equation has only one solution y such that y and y' exist for positive x and have an asymptotic expansion. By substituting (5) and (6) into (13) and equating the coefficient of each power of x to zero we obtain

$$-c_0 = 0, \quad -c_1 + 1 = 0, \quad -c_1 - c_2 = 0, \quad \dots, \quad -nc_n - c_{n+1} = 0, \quad \dots,$$

that is,

$$c_0 = 0, \quad c_1 = 1, \quad c_2 = -1, \quad \dots, \quad c_{n+1} = (-1)^n n!, \quad \dots$$

and, therefore,

$$(14) \quad f(x) \sim \frac{1}{x} - \frac{1}{x^2} + \frac{2!}{x^3} - \frac{3!}{x^4} + \dots$$

The special incomplete gamma function [cf. (12) in Sec. 17.1]

$$\text{Ei}(x) = \int_x^\infty \frac{e^{-t}}{t} dt \quad (x > 0)$$

is called the exponential integral; clearly, $\text{Ei}(x) = Q(0, x)$. From (14) we obtain the asymptotic expansion

$$(15) \quad \text{Ei}(x) = e^{-x} f(x) \sim e^{-x} \left(\frac{1}{x} - \frac{1}{x^2} + \frac{2!}{x^3} - \frac{3!}{x^4} + \dots \right).$$

Problems

Using the method illustrated by Ex. 1, find the asymptotic expansion of

1. $\text{erf } x$ [Hint: show that $y = \frac{1}{2}\sqrt{\pi} e^{x^2} \text{erfc } x$ satisfies $y' - 2xy + 1 = 0$.]

2. $Q(\frac{1}{2}, x)$ [Hint: consider $y = e^x \sqrt{x} Q(\frac{1}{2}, x)$.]

3. $Q(\alpha, x)$ [Hint: consider $y = e^x x^{1-\alpha} Q(\alpha, x)$.]

Using integration by parts, find the asymptotic expansion of

4. $\text{si}(x)$ 5. $\text{ci}(x)$ 6. $\text{c}(x)$ 7. $\text{s}(x)$ 8. $Q(0, x)$ 9. $Q(\alpha, x)$

Using the results of Probs. 4–9, find the asymptotic expansion of

10. $\text{Si}(x)$ 11. $\text{C}(x), \text{S}(x)$ 12. $P(\alpha, x)$

13. (Logarithmic integral) The function

$$\text{li}(x) = \int_0^x \frac{dt}{\ln t}$$

is called the logarithmic integral. Show that $\text{li}(e^{-x}) = -\text{Ei}(x)$.

14. Show that $\text{Ei}(ix) = \text{ci}(x) - i \text{ si}(x)$. Then replace x by ix in (15) and separate the real and imaginary parts on both sides; show that this leads to the asymptotic expansions of $\text{ci}(x)$ and $\text{si}(x)$ as obtained in Probs. 4 and 5.

15. It can be shown that if x in the asymptotic expansion of $\text{Si}(x)$ is replaced by $z = x + iy$ (z not pure imaginary), then one obtains an asymptotic expansion of the complex sine integral

$$\text{Si}(z) = \int_0^z \frac{\sin t}{t} dt.$$

Conclude from this expansion that approximations for the zeros $z_n = x_n + iy_n$ of $\text{Si}(z)$ are obtained from

$$\frac{\cos z}{z} = \frac{\pi}{2} \quad \text{or} \quad \cos z = \frac{\pi z}{2}.$$

Show that by separating the real and the imaginary parts [cf. (10), Sec. 10.6] and using $\cosh y \approx e^y/2$, $\sinh y \approx e^y/2$ (y large) we find

$$(16) \quad (a) e^y \cos x = \pi x \quad (b) e^y \sin x = -\pi y.$$

Let $x > 0$. Derive from (16)

$$\cos x = \pi x e^{-x \tan x}$$

and conclude that, since $\cos x$ is relatively small in absolute value, x_n must be close to $2n\pi$, say, $x_n = 2n\pi + \delta_n$ where δ_n is small in absolute value, so that $\sin \delta_n \approx \delta_n$, $\cos \delta_n \approx 1$, and from (16),

$$(17) \quad y_n \approx \ln(2n\pi^2), \quad x_n \approx 2n\pi - \frac{\ln(2n\pi^2)}{2n\pi}.$$

Show that from (17),

$$z_1 \approx 5.81 + 2.98i, \quad z_2 \approx 12.27 + 3.68i, \quad z_3 \approx 18.63 + 4.08i.$$

(Actually,

$$z_1 = 5.92 + 2.90i, \quad z_2 = 12.34 + 3.67i, \quad z_3 = 18.69 + 4.08i,$$

the given decimal places being correct.)

Chapter 18

Probability and Statistics

The importance of mathematical statistics in engineering is increasing, in particular in mass production and in the analysis of experimental data. In the present chapter we shall give an introduction to this field. We shall start with the tabular and graphical representation of data (Secs. 18.2, 18.3). Then we shall consider some fundamental concepts and relations from the theory of mathematical probability (Secs. 18.4–18.11), because probability theory forms the basis of mathematical statistics. The remaining sections will be devoted to some of the most important statistical methods.

Prerequisite for this chapter: differential and integral calculus.

Sections which may be omitted in a shorter course: 18.17–18.20.

References: Appendix 1, Part I.

Answers to problems: Appendix 2.

18.1 *Nature and Purpose of Mathematical Statistics*

In engineering statistics we are concerned with methods for designing and evaluating experiments to obtain information about practical problems, for example, the inspection of quality of raw material or manufactured products, the comparison of machines and tools or methods used in production, the output of workers, the reaction of consumers to new products, the yield of a chemical process under various conditions, the relation between the iron content of iron ore and the density of the ore, the effectivity of air conditioning systems under various temperatures, the relation between the Rockwell hardness and the carbon content of steel, and so on.

For instance, in a process of mass production (of screws, bolts, light bulbs, typewriter keys, etc.) we ordinarily have *nondefective items*, that is, articles which satisfy the quality requirements, and *defective items*, those which do not satisfy the requirements. Such requirements include maximum and minimum diameters of axle shafts, minimum lifetimes of light bulbs, limiting values of the resistance of resistors in radio and TV production, maximum weights of airmail envelopes, minimum contents of automatically

filled bottles, maximum reaction times of switches, and minimum values of strength of yarn.

The reason for the differences in the quality of products is *variation* due to numerous factors (in the raw material, function of automatic machinery, workmanship, etc.) whose influence cannot be predicted, so that the variation must be regarded as a *random variation*. In the case of evaluating the effectiveness of production methods and the other preceding examples the situation is similar.

In most cases the inspection of each item of a production is prohibitively expensive and time-consuming. It may even be impossible if it leads to the destruction of the item. Hence instead of inspecting all the manufactured items just a few of them (a "sample") are inspected and from this inspection conclusions can be drawn about the totality (the "population"). If we draw 100 screws from a lot of 10,000 screws and find that 5 of the 100 screws are defective, we are inclined to conclude that *about* 5% of the lot are defective, provided the screws were drawn "*at random*", that is, so that each screw of the lot had the same "chance" of being drawn. It is clear that such a conclusion is not absolutely certain, that is, we cannot say that the lot contains *precisely* 5% defectives, but in most cases such a precise statement would not be of particular practical interest anyway. We also feel that the conclusion is the more dependable the more randomly selected screws we inspect. Using the theory of mathematical probability we shall see that those somewhat vague intuitive notions can be made precise. Furthermore, this theory will also yield measures for the reliability of conclusions about populations obtained from samples by statistical methods.

Similarly, to obtain information about the iron content μ of iron ore, we may pick a certain number n of specimens at random and measure the iron content. This yields a sample of n numbers x_1, \dots, x_n (the results of the n measurements) whose average $\bar{x} = (x_1 + x_2 + \dots + x_n)/n$ is an approximate value for μ .

Problems of differing natures may require different methods and techniques, but the steps leading from the formulation to the solution of the problem are the same in almost all cases. They are as follows.

1. *Formulation of the problem.* It is important that the problem be precisely formulated and limited so that a useful answer can be expected within a prescribed interval of time, taking into account the cost of the statistical investigation, the skill of the investigators, and the facilities available. This step must also include the creation of a mathematical model based on clear concepts. (For example, we must define what we mean by a defective item in a specific situation.)

2. *Design of experiment.* This step includes the choice of the statistical method to be used in the last step, the sample size n (number of items to be drawn and inspected or number of experiments to be made, etc.), and the physical methods and techniques to be used in the experiment.

3. *Experimentation.*

4. Tabulation. In this step we arrange the experimental data in a clear and simple tabular form and represent them graphically if necessary. Furthermore, we compute numbers (called *statistics*) which characterize certain properties of the sample such as the average size and spread of the values.

5. Statistical inference. In this step we use the sample and apply a suitable statistical method for drawing conclusions about the unknown properties of the population so that we obtain the answer to our problem.

18.2 *Tabular and Graphical Representation of Samples*

In the course of a statistical experiment we usually obtain a sequence of observations (numbers in most cases) which we write down in the order they occur. A typical example is shown in Table 27. These data were obtained by making standard test cylinders (diameter 6 in., length 12 in.) of concrete and splitting them 28 days later. We thus have a **sample** consisting of 100 **sample values**, so that the **size** of the sample is $n = 100$.

Table 27. Sample of 100 Values of the Splitting Tensile Strength (lb/in.²) of Concrete Cylinders.

320	380	340	410	380	340	360	350	320	370
350	340	350	360	370	350	380	370	300	420
370	390	390	440	330	390	330	360	400	370
320	350	360	340	340	350	350	390	380	340
400	360	350	390	400	350	360	340	370	420
420	400	350	370	330	320	390	380	400	370
390	330	360	380	350	330	360	300	360	360
360	390	350	370	370	350	390	370	370	340
370	400	360	350	380	380	360	340	330	370
340	360	390	400	370	410	360	400	340	360

D. L. IVEY, Splitting tensile tests on structural lightweight aggregate concrete. Texas Transportation Institute, College Station, Texas, 1965.

In the present section we want to learn how to represent samples in a suitable tabular and graphical form. It suffices to discuss the corresponding methods in terms of our sample in Table 27.

To see what information is contained in Table 27, we shall now order the data. We write 300 (the smallest value), 310, 320, ···, 440 (the largest value) in a column. Then we run through Table 27 line by line and make a tally mark for each occurring number. In this way we obtain a **tally chart** of our sample, consisting of the first two columns of Table 28. The number of tallies is then listed in the third column of Table 28. It indicates how often the corresponding value x occurs in the sample and is called the *absolute frequency* or, more briefly; the **frequency** of that value x in the sample. Dividing it by

the size n of the sample, we obtain the **relative frequency** in column 4; in our case, $n = 100$. Thus $x = 330$ has the frequency 6 and the relative frequency 0.06 or 6%.

Table 28. Frequency Table of the Sample in Table 27.

1 Tensile Strength $x[\text{lb/in.}^2]$	2 Frequency	3 Tallies	4 Relative Frequency	5 Cumulative Frequency	6 Cumulative Relative Frequency
300		2	0.02	2	0.02
310		0	0.00	2	0.02
320		4	0.04	6	0.06
330		6	0.06	12	0.12
340		11	0.11	23	0.23
350		14	0.14	37	0.37
360		16	0.16	53	0.53
370		15	0.15	68	0.68
380		8	0.08	76	0.76
390		10	0.10	86	0.86
400		8	0.08	94	0.94
410		2	0.02	96	0.96
420		3	0.03	99	0.99
430		0	0.00	99	0.99
440		1	0.01	100	1.00

If for a certain x we sum all the frequencies corresponding to the sample values which are smaller than or equal to that x , we obtain the **cumulative frequency** corresponding to that x . This yields column 5 in Table 28. For example, to $x = 350$ there corresponds the cumulative frequency 37, and this tells us that there are 37 sample values which are smaller than or equal to 350. Division by the size n of the sample yields the **cumulative relative frequency** in column 6. For example, from column 6 we see that 76% of the sample values are smaller than or equal to 380.

If a certain numerical value does not occur in a sample, its frequency is 0. If all the n values of the sample are numerically equal, then this number has the frequency n and, therefore, the relative frequency $n/n = 1$. Since these are the two extreme possible cases, we have

Theorem 1. *The relative frequency is at least equal to 0 and at most equal to 1.*

Suppose that a given sample of size n consists of m numerically different values

$$x_1, \quad x_2, \quad \dots \quad (m \leq n)$$

with corresponding relative frequencies

$$f_1, \quad f_2, \quad \dots \quad f_m.$$

Then we may introduce the function $\tilde{f}(x)$ which for each $x = x_j$ equals the corresponding relative frequency f_j and equals 0 for every x not appearing in the sample.¹ In formulas,

$$(1) \quad \tilde{f}(x) = \begin{cases} f_j & \text{when } x = x_j, \\ 0 & \text{otherwise.} \end{cases} \quad (j = 1, 2, \dots, m)$$

This function is called the **frequency function of the sample**. It shows how the values of the sample are distributed. We therefore say that it determines the **frequency distribution** of the sample.

For example, in Table 28 the values of the frequency function are shown in column 4, and we see that $\tilde{f}(300) = 0.02$, $\tilde{f}(310) = 0$, $\tilde{f}(320) = 0.04$, and so on.

The sum of all the frequencies in a sample of size n must equal n . (Why?) This yields

Theorem 2. *The sum of all the relative frequencies in a sample equals 1, that is,*

$$\sum_{j=1}^n \tilde{f}(x_j) = \tilde{f}(x_1) + \tilde{f}(x_2) + \dots + \tilde{f}(x_m) = 1.$$

Samples may be represented in graphical form in various ways, as is illustrated by Figs. 380–383.

We now introduce the function

$\tilde{F}(x) = \text{sum of the relative frequencies of all the values that are smaller than } x \text{ or equal to } x.$

This function is called the **cumulative frequency function of the sample** or the **sample distribution function**. An example is shown in Fig. 384 on p. 738.

$\tilde{F}(x)$ is a *step function* (piecewise constant function) having jumps of magnitude $f(x)$ precisely at those x at which $\tilde{f}(x) \neq 0$. The first jump is at the smallest sample value and the last at the largest. Afterwards $\tilde{F}(x) = 1$.

The relation between $\tilde{f}(x)$ and $\tilde{F}(x)$ is

$$(2) \quad \tilde{F}(x) = \sum_{t \leq x} \tilde{f}(t)$$

where $t \leq x$ means that for an x we have to form the sum of all those $\tilde{f}(t)$ for which t is less than x or equal to x .

If a sample consists of too many numerically different sample values, then its tabular and graphical representations are unnecessarily complicated, but may be simplified by the process of **grouping**, as follows.

A sample being given, we choose an interval I that contains all the sample values. We subdivide I into subintervals, which are called **class intervals**. The midpoints of these intervals are called **class midpoints**. The sample values in each such interval are said to form a **class**. Their number is called the

¹ We write \tilde{f} because we want to reserve the simple symbol f for the theoretical counterpart of the frequency function. f will appear much more often in our text.

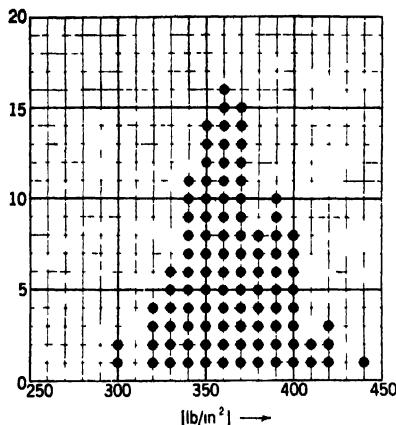


Fig. 380. Dot frequency diagram of the sample in Table 28.

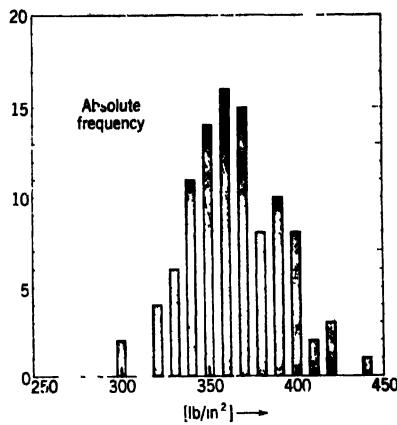


Fig. 381. Bar chart of the sample in Table 28.

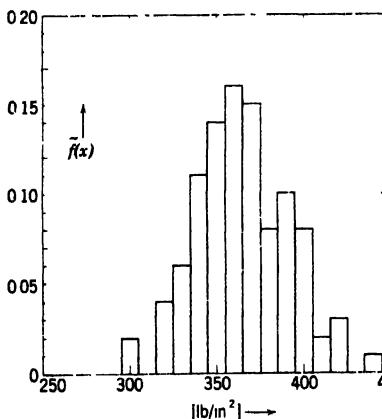


Fig. 382. Frequency histogram of the sample in Table 28.

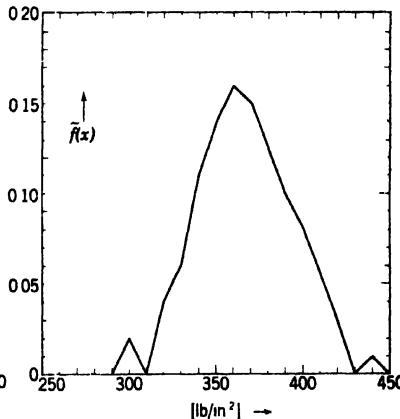


Fig. 383. Frequency polygon of the sample in Table 28.

corresponding class frequency. Division by the sample size n yields the **relative class frequency**. This frequency, considered as a function $\tilde{f}(x)$ of the class midpoints, is called the **frequency function of the grouped sample**, and the corresponding cumulative relative class frequency, considered as a function $\tilde{F}(x)$ of the class midpoints, is called the **distribution function of the grouped sample**. An example is shown in Tables 29 and 30 on p. 738.

The fewer classes we choose, the simpler the distribution of the sample becomes but the more information we lose, because the original sample values do no longer appear explicitly. Grouping should be done such that only unessential details are eliminated. Unnecessary complications in the later use of a grouped sample are avoided by obeying the following rules:

1. All class intervals should have the same length.

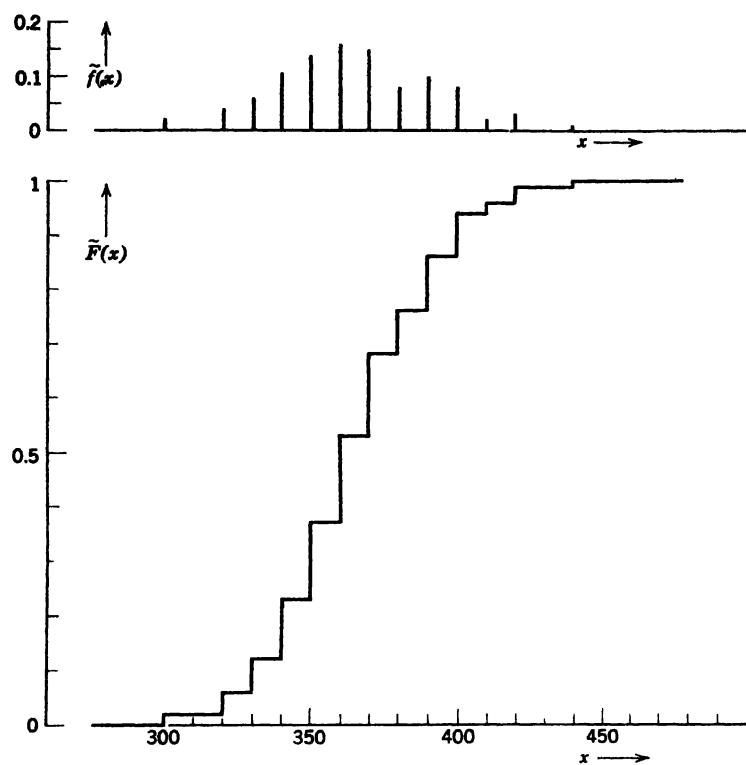


Fig. 384. Frequency function $\tilde{f}(x)$ and cumulative frequency function $\tilde{F}(x)$ of the sample in Table 28.

Table 29. Strength of 50 Lots of Cotton (lb required to break a skein).

114	118	86	107	87	94	82	81	98	84
120	126	98	89	114	83	94	106	96	111
123	110	83	118	83	96	96	74	91	81
102	107	103	80	109	71	96	91	86	129
130	104	86	121	96	96	127	94	102	87

Cotton Economic Research, The University of Texas, Austin, Texas.

Table 30. Frequency Table of the Grouped Sample in Table 29.

Class Interval	Class Midpoint x	Absolute Frequency		$\tilde{f}(x)$	$\tilde{F}(x)$
		Tallies	Sum		
65–75	70		2	0.04	0.04
75–85	80		8	0.16	0.20
85–95	90		11	0.22	0.42
95–105	100		12	0.24	0.66
105–115	110		8	0.16	0.82
115–125	120		5	0.10	0.92
125–135	130		4	0.08	1.00

2. The class midpoints should correspond to simple numbers (numbers with few figures).

3. If a sample value x , coincides with the common endpoint of two class intervals, we take it into that class interval which extends from x , to the right.

Problems

In each case make a frequency table of the given sample and represent the sample by a dot frequency diagram, a bar chart, and a histogram.

1. Carbon content [%] of coal

87 86 85 87 86 87 86 81 77 85
86 84 83 83 82 84 83 79 82 73

2. Tensile strength [kg/mm²] of sheet steel

44 43 41 41 44 44 43 44 42 45 43 43 44 45 46
42 45 41 44 44 43 44 46 41 43 45 45 45 42 44 44

3. Numbers that turned up on a die

6 2 4 1 2 4 3 3 2 1 6 5 6 3 4

4. Release time [sec] of a relay

1.3 1.4 1.1 1.5 1.4 1.3 1.2 1.4 1.5 1.3
1.2 1.3 1.5 1.4 1.4 1.6 1.3 1.5 1.1 1.4

5. Resistance [ohms] of resistors

99 100 102 101 98 103 100 102 99 101
100 100 99 101 100 102 99 101 98 100

Graph the cumulative frequency functions of the following samples

6. The sample in Prob. 3

7. The sample in Prob. 1

8. Group the sample in Table 27, using class intervals with midpoints 300, 320, 340, ... Make up the corresponding frequency table. Graph the histogram and compare it with Fig. 382. Graph the cumulative frequency function.

9. Graph the bar chart, the histogram, and the frequency polygon of the grouped sample in Table 30.

10. Group the sample in Table 29, using class intervals with midpoints 75, 85, 95, ... Make the corresponding frequency table. Graph the histogram and compare it with that in Prob. 9.

11. Graph the histogram of the following sample of lifetimes [hours] of light bulbs.

Lifetime	Absolute Frequency	Lifetime	Absolute Frequency	Lifetime	Absolute Frequency
950-1050	4	1350-1450	51	1750-1850	20
1050-1150	9	1450-1550	58	1850-1950	9
1150-1250	19	1550-1650	53	1950-2050	3
1250-1350	36	1650-1750	37	2050-2150	1

18.3 Sample Mean and Sample Variance

The frequency function (or equally well the distribution function) characterizes a given sample in detail. From this function we may compute certain numbers which measure general properties, such as the average size of the sample values, the "spread," the "asymmetry," etc. Such a constant is called a **statistic** of the sample. The two most important of these constants are the sample mean and the sample variance.

The *mean value of a sample* x_1, x_2, \dots, x_n , or, briefly, **sample mean**, is denoted by \bar{x} and is defined by the formula

$$(1) \quad \bar{x} = \frac{1}{n} \sum_{j=1}^n x_j = \frac{1}{n} (x_1 + x_2 + \dots + x_n).$$

It is the sum of all the sample values, divided by the size n of the sample. Obviously, it measures the average size of the sample values, and sometimes the term *average* is used for \bar{x} .

The *variance of a sample* x_1, x_2, \dots, x_n , or, briefly, **sample variance**, is denoted by s^2 and is defined by the formula

$$(2) \quad s^2 = \frac{1}{n-1} \cdot \sum_{j=1}^n (x_j - \bar{x})^2$$

$$\frac{1}{n-1} [(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2].$$

It is the sum of the squares of the deviations of the sample values from the mean \bar{x} , divided by $n-1$. It measures the spread or dispersion of the sample values and is positive, except for the rare case when all the sample values are equal (and are then equal to \bar{x}). The positive square root of the sample variance s^2 is called the **standard deviation** of the sample and is denoted by s .

Example 1. Ten randomly selected nails had the lengths (in.)

0.80 0.81 0.81 0.82 0.81 0.82 0.80 0.82 0.81 0.81

From (1) we see that the sample mean is

$$\bar{x} = \frac{1}{10}(0.80 + 0.81 + 0.81 + 0.82 + \dots + 0.81) = 0.811 \text{ [in.]}$$

Applying (2), we thus obtain the sample variance

$$s^2 = \frac{1}{9}[(0.800 - 0.811)^2 + \dots + (0.810 - 0.811)^2] = 0.000\,054 \text{ [in.}^2\text{].}$$

This calculation becomes simpler if we take equal sample values together. Then

$$\bar{x} = \frac{1}{10}(2 \cdot 0.80 + 5 \cdot 0.81 + 3 \cdot 0.82) = 0.811.$$

In the parentheses we have the sum of the three *numerically different* sample values $x_1 = 0.80$, $x_2 = 0.81$, $x_3 = 0.82$, each multiplied by its frequency. Similarly,

$$s^2 = \frac{1}{9} [2(0.800 - 0.811)^2 + 5(0.810 - 0.811)^2 + 3(0.820 - 0.811)^2] = 0.000\,054.$$

The example illustrates how we may compute \bar{x} and s^2 by the use of the frequency function $\tilde{f}(x)$ of the sample. If a sample of n values contains precisely m *numerically different* sample values

$$x_1, \quad x_2, \quad \dots, \quad x_m \quad (m \leq n),$$

the corresponding relative frequencies are

$$\tilde{f}(x_1), \quad \tilde{f}(x_2), \quad \dots, \quad \tilde{f}(x_m).$$

Hence the corresponding frequencies needed in the computation are

$$n\tilde{f}(x_1), \quad n\tilde{f}(x_2), \quad \dots, \quad n\tilde{f}(x_m),$$

and (1) and (2) take the form

$$(3) \quad \bar{x} = \frac{1}{n} \sum_{j=1}^m x_j n\tilde{f}(x_j)$$

and

$$(4) \quad s^2 = \frac{1}{n-1} \sum_{j=1}^m (x_j - \bar{x})^2 n\tilde{f}(x_j).$$

Note that in (1) and (2) we sum over *all* the sample values, whereas now we sum over the *numerically different* sample values. The absolute frequencies $n\tilde{f}(x_j)$ are integers, while the relative frequencies $\tilde{f}(x_j)$ may be messy, for example, if $n = 23$ or $n = 84$, etc.

Our formulas for s^2 are impractical because \bar{x} and therefore $x_j - \bar{x}$ will in general have to be rounded and thus suffer from rounding errors. We shall now derive the formula for s^2 which is used in computations. Inserting

$$(x_j - \bar{x})^2 = x_j^2 - 2x_j\bar{x} + \bar{x}^2$$

into (2) and decomposing into three sums, we have

$$\sum (x_j - \bar{x})^2 = \sum x_j^2 - 2\bar{x} \sum x_j + \sum \bar{x}^2.$$

The last sum equals $n\bar{x}^2$. Replacing \bar{x} according to (1), we obtain

$$-2\bar{x} \sum x_j = -\frac{2}{n} (\sum x_j)^2 \quad \text{and} \quad n\bar{x}^2 = \frac{1}{n} (\sum x_j)^2.$$

Taking these two expressions together, we finally have

$$(5) \quad s^2 = \frac{1}{n-1} \left[\sum_{j=1}^n x_j^2 - \frac{1}{n} \left(\sum_{j=1}^n x_j \right)^2 \right].$$

Similarly, (4) may be transformed to

$$(6) \quad s^2 = \frac{1}{n-1} \left[\sum_{j=1}^m x_j^2 n\tilde{f}(x_j) - \frac{1}{n} \left(\sum_{j=1}^m x_j n\tilde{f}(x_j) \right)^2 \right].$$

For instance, in Ex. 1 we obtain from (3) and (6) (cf. Table 31) the values $\bar{x} = 8.11/10 = 0.811$ and

$$s^2 = \frac{1}{9} \left(6.5777 - \frac{8.11^2}{10} \right) = \frac{0.00049}{9} = 0.000054,$$

as before.

Table 31. Computation of Mean and Variance in Ex. 1 by the Use of (3) and (6).

x_j	$10\tilde{f}(x_j)$	$x_j \cdot 10\tilde{f}(x_j)$	x_j^2	$x_j^2 \cdot 10\tilde{f}(x_j)$
0.80	2	1.60	0.6400	1.2800
0.81	5	4.05	0.6561	3.2805
0.82	3	2.46	0.6724	2.0172
Sum		8.11		6.5777

Often a further important simplification is obtained by setting

$$(7) \quad x_j = x_j^* + c \quad (x_j^* = x_j - c)$$

where the constant c is chosen such that the transformed values x_j^* are small in absolute value. Geometrically this means that we shift the origin on the scale, and this procedure is sometimes called the *method of working origin*. We now compute the mean \bar{x}^* of the transformed sample and then the mean

$$(8) \quad \bar{x} = c + \bar{x}^* \quad \text{where} \quad \bar{x}^* = \frac{1}{n} \sum_{j=1}^n x_j^*.$$

Furthermore, the variance of the original sample is the same as the variance s^{*2} of the transformed sample, that is,

$$(9) \quad s^2 = s^{*2} \quad \text{where} \quad s^{*2} = \frac{1}{n-1} \left[\sum_{j=1}^n x_j^{*2} - \frac{1}{n} \left(\sum_{j=1}^n x_j^* \right)^2 \right]$$

The proof is left to the reader (Prob. 11). Table 32 shows an example, clearly, $\tilde{f}(x_j^*) = \tilde{f}(x_j)$.

Table 32. Computation of Mean \bar{x} and Variance s^2 in Ex. 1 by Means of (8) and (9) with $c = 0.80$.

x_j	$x_j^* = x_j - 0.80$	$10\tilde{f}(x_j^*)$	$x_j^* \cdot 10\tilde{f}(x_j^*)$	x_j^{*2}	$x_j^{*2} \cdot 10\tilde{f}(x_j^*)$
0.80	0.00	2	0.00	0.0000	0.0000
0.81	0.01	5	0.05	0.0001	0.0005
0.82	0.02	3	0.06	0.0004	0.0012
Sum			0.11		0.0017

$$\bar{x} = 0.800 + \bar{x}^* = 0.800 + 0.011 = 0.811$$

$$s^{*2} = \frac{1}{9} \left(0.0017 - \frac{0.11^2}{10} \right) = 0.000054$$

The method of working origin is a particular case of the **method of full coding**. In the latter method we set

$$(10) \quad x_j = c_1 x_j^* + c_2, \quad \left[\text{thus } x_j^* = \frac{1}{c_1} (x_j - c_2) \right]$$

with conveniently chosen constants c_1 and c_2 . We first compute the mean \bar{x}^* and the variance s^{*2} of the transformed sample and then

$$(11) \quad \bar{x} = c_1 \bar{x}^* + c_2$$

and

$$(12) \quad s^2 = c_1^2 s^{*2}.$$

Examples are shown in Tables 33 and 34; we see that in Table 34 formula (10) gives $x_j^* = 0.1(x_j - 350)$.

Table 33. Computation of Mean \bar{x} and Variance s^2 in Ex. 1 by Means of (10), (11), and (12) with $c_1 = 0.01$ and $c_2 = 0.80$.

x_j	$x_j^* = 100(x_j - 0.80)$	$10\tilde{f}(x_j^*)$	$x_j^* \cdot 10\tilde{f}(x_j^*)$	x_j^{*2}	$x_j^{*2} \cdot 10\tilde{f}(x_j^*)$
0.80	0	2	0	0	0
0.81	1	5	5	1	5
0.82	2	3	6	4	12
Sum			11		17

$$\bar{x} = 0.01 \cdot 11 + 0.80 = 0.811, \quad s^2 = 0.0001 \cdot s^{*2} = \frac{0.0001}{9} \left(17 - \frac{11^2}{10} \right) = 0.000054$$

Table 34. Computation of Mean \bar{x} and Variance s^2 of the Sample in Table 28 by means of (11) and (12) with $c_1 = 10$ and $c_2 = 350$.

x_j	x_j^*	$100\tilde{f}(x_j^*)$	$x_j^* \cdot 100\tilde{f}(x_j^*)$	x_j^{*2}	$x_j^{*2} \cdot 100\tilde{f}(x_j^*)$
300	-5	2	-10	25	50
310	-4	0	0	16	0
320	-3	4	-12	9	36
330	-2	6	-12	4	24
340	-1	11	-11	1	11
350	0	14	0	0	0
360	1	16	16	1	16
370	2	15	30	4	60
380	3	8	24	9	72
390	4	10	40	16	160
400	5	8	40	25	200
410	6	2	12	36	72
420	7	3	21	49	147
430	8	0	0	64	0
440	9	1	9	81	81
Sum		100	147		929

$$\bar{x}^* = \frac{147}{100} = 1.47 \quad s^{*2} = \frac{1}{99} \left(929 - \frac{147^2}{100} \right) = 7.201$$

$$\bar{x} = 10 \cdot 1.47 + 350 = 364.7 \quad s^2 = 10^2 \cdot 7.201 = 720.1$$

Problems

Compute the mean and the variance of the following samples.

1. The sample in Prob. 1, Sec. 18.2
2. The sample in Prob. 2, Sec. 18.2
3. The sample in Prob. 3, Sec. 18.2
4. The sample in Prob. 11, Sec. 18.2
5. Compute the mean and the variance of the ungrouped sample in Table 29 and the grouped sample in Table 30 and compare the results.
6. If a sample is being grouped, its mean will change, in general. Show that the change cannot exceed $l/2$ where l is the length of each class interval.
7. The p th percentile of a sample is a number Q_p such that at least $p\%$ of the sample values are smaller than or equal to Q_p , and also at least $(100 - p)\%$ of those values are larger than or equal to Q_p . If there is more than one such number (in which case there will be an interval of them), the p th percentile is defined as the average of the numbers (midpoint of the interval). In particular, Q_{50} is called the **middle quartile** or **median** and is denoted by \bar{x} . Find \bar{x} for the sample in Table 28.
8. The percentiles Q_{25} and Q_{75} of a sample are called the **lower** and **upper quartiles** of the sample, and $Q_{75} - Q_{25}$, which is a measure for the spread, is called the **interquartile range**. Find Q_{25} , Q_{75} , and $Q_{75} - Q_{25}$ for the sample in Table 28.
9. Same tasks as in Probs. 7 and 8 for the sample in Table 29.
10. Derive (6) from (4).
11. Derive (8) and (9).
12. Derive (11) and (12).

18.4 Random Experiments, Outcomes, Events

Statistical experiments or observations yield samples from which we want to draw conclusions about the corresponding population. Before we can do so we must first develop mathematical models of populations, using the theory of mathematical probability. This theory, therefore, supplies an important basis in mathematical statistics, and we shall consider it to the extent needed. In this section we shall introduce some fundamental concepts.

A *random experiment* or *random observation*, briefly **experiment** or **observation**, is a process that has the following properties.

1. It is performed according to a set of rules that determines the performance completely.
2. It can be repeated arbitrarily often.
3. The result of each performance depends on "chance" (that is, on influences which we cannot control) and can therefore not be uniquely predicted.

The result of a single performance of the experiment is called the **outcome** of that trial.

Examples are games of chance such as rolling a die, flipping a coin, or drawing two cards from a shuffled deck of playing cards, technical experiments such as the random selection and inspection of 10 screws from a box

containing 100 screws, the random selection of 5 pieces of sheet steel and the measurement of the Rockwell hardness, the measurement of the lifetimes of 5 light bulbs randomly selected from a lot of bulbs, the determination of the yield of a chemical process under various conditions, the relation between the working conditions (temperature, humidity, etc.) and the output, the random selection of 20 persons from a group and the determination of their blood pressures or their opinions about a certain movie, etc.

The set of all possible outcomes of an experiment is called the **sample space** of the experiment and will be denoted by S . Each outcome is called an *element* or *point* of S . A sample space is said to be *finite* or *infinite*, depending on whether it consists of finitely many or infinitely many elements, respectively.

For example, if we roll a die once, the die can turn up one of the 6 faces numbered 1, 2, ⋯, 6, and the sample space consists of 6 elements, so that we may write

$$S = \{1, 2, 3, 4, 5, 6\}.$$

In connection with industrial production we may draw an item to find out whether it is defective or nondefective. Then S consists of the two elements D (defective) and N (nondefective), which may also be characterized by numbers, for example 0 (defective) and 1 (nondefective). If in this case we distinguish between several types of defects, we have a sample space consisting of more than two points.

The sample space of the experiment of measuring the strength of cotton (cf. Table 29 in Sec. 18.2) is infinite, because the outcome may be any positive number within a certain range.

In most practical problems we are not so much interested in the individual outcomes but in whether an outcome belongs (or does not belong) to a certain set of outcomes. Clearly, each such set A is a subset of the sample space S . It is called an **event**.

Example 1. If we draw 2 gaskets from a set of 5 gaskets (numbered from 1 to 5), the sample space consists of the 10 possible outcomes

$$1, 2 \quad 1, 3 \quad 1, 4 \quad 1, 5 \quad 2, 3 \quad 2, 4 \quad 2, 5 \quad 3, 4 \quad 3, 5 \quad 4, 5, .$$

but we may be interested in the number of defective gaskets we get in the draws and thus distinguish among the 3 events

$$A: \text{No defective gasket}, \quad B: 1 \text{ defective gasket}, \quad C: 2 \text{ defective gaskets}.$$

Assuming that 3 gaskets, say 1, 2, 3, are defective, we see that

$$A \text{ occurs, if we draw } 4, 5$$

$$B \text{ occurs, if we draw } 1, 4; \quad 1, 5; \quad 2, 4; \quad 2, 5; \quad 3, 4; \quad \text{or } 3, 5$$

$$C \text{ occurs, if we draw } 1, 2; \quad 1, 3; \quad \text{or } 2, 3.$$

A sample space S and the events of an experiment can be represented graphically by a **Venn diagram**, as follows. Suppose that the set of points inside the rectangle in Fig. 385 represents S . Then the interior of a closed curve inside the rectangle represents an event which we denote by E . The

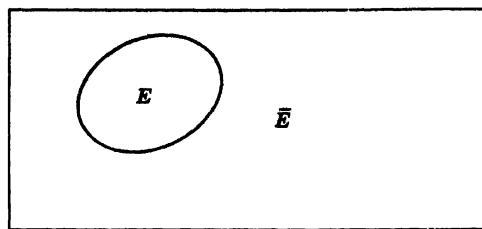


Fig. 385. Venn diagram representing a sample space S and events E and \bar{E} .

set of all the elements (outcomes) not in E is called the **complement of E** in S and is denoted by \bar{E} .

For example, in rolling a die once the complement of the event

E : *The die turns up an even number*

is the event

\bar{E} : *The die turns up an odd number.*

An event containing no element is called the **impossible event** or *empty event* and is denoted by ϕ .

The notion of event includes each outcome as well as the entire sample space S as special cases.

Let A and B be any two events in an experiment. Then the event consisting of all the elements of the sample space S contained in A or B , or both, is called the **union** of A and B and is denoted by

$$A \cup B.$$

The event consisting of all the elements in S contained in both A and B is called the **intersection** of A and B and is denoted by

$$A \cap B.$$

Figure 386 illustrates how to represent these two events by a Venn diagram. If A and B have no element in common, then $A \cap B = \phi$, and A and B are called **mutually exclusive events** or *disjoint events*.

For instance, in Ex. 1, $B \cap C = \phi$ and

$B \cup C$: *1 or 2 defective gaskets.*

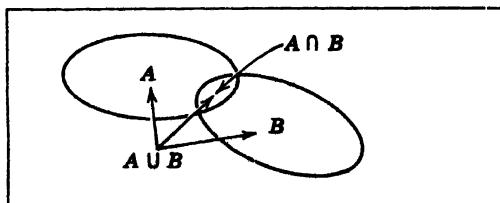


Fig. 386. Venn diagram representing two events A and B in a sample space S and their union $A \cup B$ and intersection $A \cap B$.

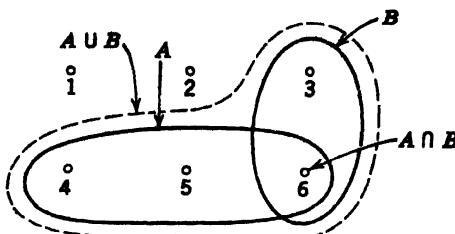


Fig. 387. Venn diagram in Example 2.

Example 2. In rolling a die once the events

- A: The die turns up a number not smaller than 4
- B: The die turns up a number divisible by 3

have the union $A \cup B = \{3, 4, 5, 6\}$ and the intersection $A \cap B = \{6\}$ (cf. Fig. 387).

If all the elements of an event A are also contained in an event B , then A is called a *subevent* of B , and we write

$$A \subset B \quad \text{or} \quad B \supset A.$$

Clearly, if $A \subset B$, then if B occurs, A necessarily occurs. For instance, the event $D = \{4; 6\}$ is a subevent of A in Ex. 2.

Let A_1, \dots, A_m be several events in a sample space S . Then the event consisting of all elements contained in one or more of these m events is called the *union* of A_1, \dots, A_m and is denoted by

$$A_1 \cup A_2 \cup \dots \cup A_m, \quad \text{or more briefly} \quad \bigcup_{j=1}^m A_j,$$

The event consisting of all elements contained in all those m events is called the *intersection* of A_1, \dots, A_m and is denoted by

$$A_1 \cap A_2 \cap \dots \cap A_m, \quad \text{or more briefly} \quad \bigcap_{j=1}^m A_j.$$

More generally, let $A_1, A_2, \dots, A_m, \dots$ be infinitely many elements in S . Then the *union*

$$A_1 \cup A_2 \cup \dots, \quad \text{or more briefly} \quad \bigcup_{j=1}^{\infty} A_j,$$

is defined to be the event consisting of all the elements contained in at least one of those events, and the *intersection*

$$A_1 \cap A_2 \cap \dots, \quad \text{or more briefly} \quad \bigcap_{j=1}^{\infty} A_j,$$

is defined to be the event consisting of all the elements contained in all those events.

If events $A_1, A_2, \dots, A_m, \dots$ are such that the occurrence of any of them makes the simultaneous occurrence of any other of them impossible, then

$A_j \cap A_k = \emptyset$ for any j and $k \neq j$, and the events are called **mutually exclusive events or disjoint events**.

For instance, the events A , B , and C in Ex. 1 are mutually exclusive.

Suppose that we perform a random experiment n times and obtain a sample consisting of n values. Let A and B be events whose relative frequencies in those n trials are $\tilde{f}(A)$ and $\tilde{f}(B)$, respectively. Then the event $A \cup B$ has the relative frequency

$$(1) \quad \tilde{f}(A \cup B) = \tilde{f}(A) + \tilde{f}(B) - \tilde{f}(A \cap B).$$

If A and B are mutually exclusive, then $\tilde{f}(A \cap B) = 0$, and

$$(2) \quad \tilde{f}(A \cup B) = \tilde{f}(A) + \tilde{f}(B).$$

These formulas are rather obvious from the Venn diagram in Fig. 386; the formal proof is left to the reader (Prob. 1).

Problems

1. Prove (1).
2. A box of 20 fountain pens contains 10 nondefective pens, 8 with type A defects, 5 with type B defects, and 3 with both types of defects. Suppose that 1 pen is drawn at random. Graph a Venn diagram of the corresponding sample space S showing the events E_A of getting a type A defective, E_B of getting a type B defective, $E_A \cap E_B$, $E_A \cup E_B$, $E_{\bar{A}} \cap E_B$, $E_{\bar{A}} \cup E_B$, $E_{\bar{A}} \cap E_{\bar{B}}$, $E_{\bar{A}} \cup E_{\bar{B}}$, and $E_A \cup E_{\bar{B}}$. How many elements (outcomes) are in each event?

18.5 Probability

Experience shows that if we perform a random experiment very often, then the occurrence of events obeys certain laws. In particular the relative frequency of an event tends to become a constant. By this we mean that if we perform several long sequences of the experiment, the corresponding relative frequencies will be almost the same. We say that the experiment shows *statistical regularity or stability of the relative frequencies*. This is illustrated by Table 35, which needs no further explanation, and Fig. 388. In Fig. 388 the point is that with increasing number of births the percentage of boys seems to

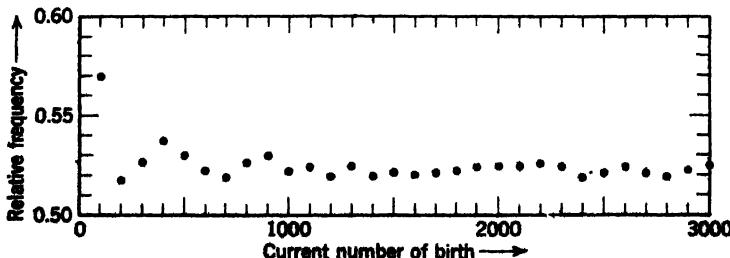


Fig. 388. Relative frequency of the event "birth of a boy." (Data from Graz, Austria, 1962.)

fluctuate less and less. The number of defective items in a production under reasonably constant conditions will show a similar behavior, and other examples can easily be found.

Table 35. Coin Tossing.

Experiments by	Number of Throws	Number of Heads	Relative Frequency of Heads
BUFFON	4,040	2,048	0.5069
K. PEARSON	12,000	6,019	0.5016
K. PEARSON	24,000	12,012	0.5005

The great majority of random experiments having practical importance exhibit this stability. We may therefore conjecture that it will be practically certain that the relative frequency of an event E in a great number of performances of such a random experiment is approximately equal to a certain number $P(E)$.

For this reason we now postulate the existence of a number $P(E)$ which is called the probability of that event E in that random experiment. Note that this number is not an absolute property of E but refers to a certain sample space S , that is, to a certain random experiment.

The statement " E has the probability $P(E)$ " then means that if we perform the experiment very often, it is practically certain that the relative frequency $\hat{f}(E)$ is approximately equal to $P(E)$. (Here the term "approximately equal" must be made precise. This we have to postpone until Sec. 18.10.)

The probability thus introduced is the counterpart of the empirical relative frequency. It is therefore natural to require that it should have certain basic properties which the relative frequency has. These properties are suggested by Theorems 1 and 2, Sec. 18.2, and formula (2), Sec. 18.4, and may be formulated as so-called

Axioms of mathematical probability

1. If E is any event in a sample space S , then

$$(1) \quad 0 \leq P(E) \leq 1.$$

2. To the entire sample space S there corresponds

$$(2) \quad P(S) = 1.$$

3. If A and B are mutually exclusive events (cf. Sec. 18.4), then

$$(3) \quad P(A \cup B) = P(A) + P(B).$$

If the sample space is infinite, we must replace Axiom 3 by

- 3*. If E_1, E_2, \dots are mutually exclusive events, then

$$(3^*) \quad P(E_1 \cup E_2 \cup \dots) = P(E_1) + P(E_2) + \dots.$$

From Axiom 3 we obtain by induction the following

Theorem 1 (Addition rule for mutually exclusive events). *If E_1, \dots, E_m are mutually exclusive events, then*

$$(4) \quad P(E_1 \cup E_2 \cup \dots \cup E_m) = P(E_1) + P(E_2) + \dots + P(E_m).$$

The reader may prove the following analogue of (1), Sec. 18.4.

Theorem 2 (Addition rule for arbitrary events). *If A and B are any events in a sample space S , then*

$$(5) \quad P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Furthermore, an event E and its complement \bar{E} (Sec. 18.4) are mutually exclusive, and $E \cup \bar{E} = S$. Using Axioms 3 and 2, we thus have

$$P(E \cup \bar{E}) = P(E) + P(\bar{E}) = 1.$$

This yields

Theorem 3 (Complementation rule). *The probabilities of an event E and its complement \bar{E} in a sample space S are related by the formula*

$$(6) \quad P(E) = 1 - P(\bar{E}).$$

This formula may be used if the computation of $P(\bar{E})$ is simpler than that of $P(E)$. An application will be shown in Ex. 2.

How can we assign probabilities to the various events in a sample space S ?

If S is finite, consisting of k elements, and the nature of the experiment shows that these k outcomes are equally likely to occur, then we may assign the same probability to each outcome, and because of Axiom 2 this probability must equal $1/k$. In this case the computation of probabilities of events reduces to that of counting the elements that make up the events.

Example 1. We consider a fair die, that is, a die of homogeneous material and strictly cubical form, and roll it once. In this experiment, $S = \{1, 2, 3, 4, 5, 6\}$. Thus $P(1) = 1/6$, $P(2) = 1/6, \dots, P(6) = 1/6$. From this and Theorem 1 we see that, for instance, the event

A: An even number turns up

has the probability $P(A) = P(2) + P(4) + P(6) = 1/2$, the event

B: A number greater than 4 turns up

has the probability $P(B) = P(5) + P(6) = 1/3$, etc. Formulas for more complicated cases and further examples will be included in the next section.

Example 2. 5 coins are tossed simultaneously. Find the probability of the event *A: At least 1 head turns up*. Since each coin can turn up heads or tails, the sample space consists of $2^5 = 32$ elements. Assuming that the coins are "true," we may assign the same probability ($1/32$) to each outcome. Then the event *A* (*No head turns up*) consists of only 1 outcome. Hence $P(\bar{A}) = 1/32$, and the answer is $P(A) = 1 - P(\bar{A}) = 31/32$.

If the nature of the experiment does not show that the finitely many outcomes are equally likely to occur, or if the sample space is not finite, we must assign

probabilities by using relative frequencies observed in long sequences of trials in such a fashion that the axioms of probability are satisfied.

In this way we only obtain approximate values, but this does not matter. It resembles the situation in classical physics, where we postulate that each body has a certain mass but cannot determine that mass exactly. This is no handicap for developing the theory.

If we are in doubt whether we assigned the probabilities in the right fashion, we may apply a certain statistical test which we shall consider later.

Often it is required to find the probability of an event B if it is known that an event A has occurred. This probability is called the **conditional probability of B given A** and is denoted by $P(B | A)$. In this case A serves as a new (reduced) sample space, and that probability is the fraction of $P(A)$ which corresponds to $A \cap B$. Thus

$$(7) \quad P(B | A) = \frac{P(A \cap B)}{P(A)} \quad [P(A) \neq 0].$$

Similarly, the **conditional probability of A given B** is

$$(8) \quad P(A | B) = \frac{P(A \cap B)}{P(B)} \quad [P(B) \neq 0].$$

Solving (7) and (8) for $P(A \cap B)$, we obtain

Theorem 4 (Multiplication rule). *If A and B are events in a sample space S and $P(A) \neq 0$, $P(B) \neq 0$, then*

$$(9) \quad P(A \cap B) = P(A)P(B | A) = P(B)P(A | B).$$

If the events A and B are such that

$$(10) \quad P(A \cap B) = P(A)P(B),$$

they are called **independent events**. Assuming $P(A) \neq 0$, $P(B) \neq 0$, we see from (7)–(9) that in this case

$$P(A | B) = P(A), \quad P(B | A) = P(B),$$

which means that the probability of A does not depend on the occurrence or nonoccurrence of B , and conversely. This justifies the terminology.

Similarly, m events A_1, \dots, A_m are said to be **independent** if

$$(10) \quad P(A_1 \cap A_2 \cap \cdots \cap A_m) = P(A_1)P(A_2) \cdots P(A_m).$$

To prepare for an illustrative example, we first note that there are two ways of drawing objects for obtaining a sample from a given set of objects, briefly referred to as **sampling from a population**, as follows.

1. **Sampling with replacement** means that the object which was drawn at random is placed back to the given set and the set is mixed thoroughly. Then we draw the next object at random.

2. Sampling without replacement means that the object which was drawn is put aside.

Example 3 (Sampling with and without replacement). A box contains 10 screws, 3 of which are defective. Two screws are drawn at random. Find the probability that none of the two screws is defective. We consider the events

A: First drawn screw nondefective

B: Second drawn screw nondefective.

Clearly, $P(A) = \frac{7}{10}$ because 7 of the 10 screws are nondefective and we sample at random so that each screw has the same probability ($\frac{1}{10}$) of being picked. If we sample with replacement, the situation before the second drawing is the same as at the beginning, and $P(B) = \frac{7}{10}$. The events are independent, and the answer is

$$P(A \cap B) = P(A)P(B) = 0.7 \cdot 0.7 = 0.49 = 49\%.$$

If we sample without replacement, then $P(A) = \frac{7}{10}$, as before. If *A* has occurred, then there are 9 screws left in the box, 3 of which are defective. Thus $P(B | A) = \frac{6}{9} = \frac{2}{3}$, and Theorem 4 yields the answer

$$P(A \cap B) = \frac{7}{10} \cdot \frac{2}{3} \approx 47\%.$$

Problems

1. Prove Theorem 1.
2. Prove Theorem 2.
3. Extending Theorem 4, show that

$$P(A \cap B \cap C) = P(A)P(B | A)P(C | A \cap B).$$

4. Find the probability of obtaining at least 1 head if 10 coins are tossed.
5. A box of 100 gaskets contains 10 gaskets with type *A* defects, 5 gaskets with type *B* defects, and 2 gaskets with both types of defects. Suppose that a gasket drawn is known to have a type *A* defect. What is the probability that it also has a type *B* defect?
6. Ten screws are drawn at random from a lot of 1000 screws, 100 of which are defective. Find the probability of the event that all 10 screws drawn are nondefective.

18.6 Permutations and Combinations

From the last section we know that in the case of a finite sample space *S* consisting of *k* equally likely outcomes, each outcome has probability $1/k$, and the probability of an event *A* is obtained by counting the number of outcomes of which *A* consists. Clearly, if that number is *m*, then $P(A) = m/k$. For the counting of outcomes the subsequent formulas will often be useful.

Suppose that there are given *n* things (*elements or objects*). We may lay them out in a row *in any order*. Each such arrangement is called a **permutation** of those things.

Theorem 1. *The number of permutations of *n* different things taken all at a time is*

$$n! = 1 \cdot 2 \cdot 3 \cdots n$$

(read "*n factorial*").



In fact, there are n possibilities for filling the first place in the row; then $n - 1$ objects are still available for filling the second place, etc. In a similar fashion we obtain

Theorem 2. *If n given things can be divided into c classes such that things belonging to the same class are alike while things belonging to different classes are different, then the number of permutations of these things taken all at a time is*

$$(2) \quad \frac{n!}{n_1! n_2! \cdots n_c!} \quad (n_1 + n_2 + \cdots + n_c = n)$$

where n_j is the number of things in the j th class.

A permutation of n things taken k at a time is a permutation containing only k of the n given things. Two such permutations consisting of the same k elements, in a different order, are different, by definition. For example, there are 6 different permutations of the 3 letters a, b, c , taken 2 letters at a time, namely ab, ac, bc, ba, ca, cb .

A permutation of n things taken k at a time with repetitions is an arrangement obtained by putting any given thing in the first position, any given thing, including a repetition of the one just used, in the second, and continuing until k positions are filled. For example, there are $3^2 = 9$ different such permutations of a, b, c , taken 2 letters at a time, namely, the preceding 6 permutations and aa, bb, cc . The reader may prove the following theorem (cf. Prob. 1).

Theorem 3. *The number of different permutations of n different things taken k at a time without repetitions is*

$$(3a) \quad n(n - 1)(n - 2) \cdots (n - k + 1) = \frac{n!}{(n - k)!}$$

and with repetitions is

$$(3b) \quad n^k.$$

In the case of a permutation not only are the things themselves important, but so also is their order. A combination of given things is any selection of one or more of the things *without regard to order*. There are two types of combinations, as follows.

The number of combinations of n different things, k at a time, *without repetitions* is the number of sets that can be made up from the n things, each set containing k different things and no two sets containing exactly the same k things.

The number of combinations of n different things, k at a time, *with repetitions* is the number of sets that can be made up of k things chosen from the given n , each being used as often as desired.

For example, there are 3 combinations of the 3 letters a, b, c , taken 2 letters at a time, without repetitions, namely, ab, ac, bc , and 6 such combinations with repetitions, namely, $ab, ac, bc, aa, bb, cc..$

Theorem 4. *The number of different combinations of n different things, k at a time, without repetitions, is*

$$(4a) \quad \binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{1\cdot2\cdots k},$$

and the number of those combinations with repetitions is

$$(4b) \quad \binom{n+k-1}{k}.$$

The statement involving (4a) follows from the first part of Theorem 3 by noting that there are $k!$ permutations of k things from the given n things which differ by the order of the elements (cf. Theorem 1), but there is only a single combination of those k things of the type characterized in the first statement of Theorem 4. The last statement of Theorem 4 can be proved by induction (cf. Prob. 2).

Example 1. If there are 10 different screws in a box which are needed in a certain order for assembling a certain product, and these screws are drawn at random from the box, the probability P of picking them in the desired order is very small, namely (cf. Theorem 1),

$$P = 1/10! = 1/3628\,800 \approx 0.000\,03\%.$$

If the box contains 6 alike right-handed screws and 4 alike left-handed screws and the 6 right-handed ones are needed first and the 4 left-handed ones are needed afterwards, the probability P that random drawing yields the screws in the desired order is (cf. Theorem 2)

$$P = 6! \cdot 4!/10! = 1/210 \approx 0.5\%.$$

Example 2. In a coded telegram the letters are arranged in groups of 5 letters, called words. From (3b) we see that there are

$$26^5 = 11\,881\,376$$

different such words. From (3a) it follows that there are

$$26!/(26-5)! = 26 \cdot 25 \cdot 24 \cdot 23 \cdot 22 = 7\,893\,600$$

different such words which contain each letter at most once.

Example 3. The number of samples of 5 light bulbs which can be selected from a lot of 500 bulbs is [cf. (4a)]

$$\binom{500}{5} = \frac{500!}{5! \cdot 495!} = \frac{500 \cdot 499 \cdot 498 \cdot 497 \cdot 496}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 255\,244\,687\,600.$$

Let us include some remarks about the *factorial function*. We define

$$(5) \quad 0! = 1$$

and may compute further values by the relation

$$(6) \quad (n+1)! = (n+1)n!.$$

For large n the function is very large (cf. Table 36). A convenient approximation for large n is the **Stirling formula**

$$(7) \quad n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \quad (e = 2.718 \cdots),$$

where \sim means that the ratio of the two sides of (7) approaches 1 as n approaches infinity.

The binomial coefficients are defined by the formula

$$(8) \quad \binom{a}{k} = \frac{a(a-1)(a-2)\cdots(a-k+1)}{k!} \quad (k \geq 0, \text{ integral}).$$

The numerator has k factors. Furthermore, we define

$$(9) \quad \binom{a}{0} = 1, \quad \text{in particular} \quad \binom{0}{0} = 1$$

For integral $a = n$ we obtain from (8)

$$(10) \quad \binom{n}{k} = \binom{n}{n-k} \quad (n \geq 0, 0 \leq k \leq n).$$

Binomial coefficients may be computed recursively, because

$$(11) \quad \binom{a}{k} + \binom{a}{k+1} = \binom{a+1}{k+1} \quad (k \geq 0, \text{ integral}).$$

Formula (8) also yields

$$(12) \quad \binom{-m}{k} = (-1)^k \binom{m+k-1}{k} \quad (k \geq 0, \text{ integral}, m > 0).$$

There are numerous further relations; we mention

$$(13) \quad \sum_{s=0}^{n-1} \binom{k+s}{k} = \binom{n+k}{k+1} \quad (k \geq 0, n \geq 1, \text{ both integral})$$

and

$$(14) \quad \sum_{k=0}^r \binom{p}{k} \binom{q}{r-k} = \binom{p+q}{r}$$

Table 36. Factorial Function.

n	$n!$	$\log(n!)$	n	$n!$	$\log(n!)$	n	$n!$	$\log(n!)$
1	1	0.000 000	6	720	2.857 332	11	39 916 800	7.601 156
2	2	0.301 030	7	5 040	3.702 431	12	479 001 600	8.680 337
3	6	0.778 151	8	40 320	4.605 521	13	6 227 020 800	9.794 280
4	24	1.380 211	9	362 880	5.559 763	14	87 178 291 200	10.940 408
5	120	2.079 181	10	3 628 800	6.559 763	15	1 307 674 368 000	12.116 500

Problems

1. Prove Theorem 3.
2. Prove the last statement of Theorem 4. Hint: use (13).
3. Using (7), compute approximate values of $4!$ and $8!$ and determine the absolute and relative errors.

4. Derive (11) from (8).
5. Find the number of samples of 10 objects which can be drawn from a lot of 100 objects.
6. How many different license plates showing 5 symbols, namely, 2 letters followed by 3 digits, could be made?
7. An urn contains 2 black, 3 white, and 4 red balls. We draw 1 ball at random and put it aside. Then we draw the next ball, and so on. Find the probability of drawing at first the two black balls, then the 3 white ones, and finally the red ones.
8. Suppose that 3 letters addressed to 3 different persons are inserted into 3 addressed envelopes at random. What is the probability that at least 1 letter is inserted into its own envelope?
9. List all permutations of 4 letters a, b, c, d taken all at a time.
10. List all permutations of the five numbers 1, 2, 3, 4, 5 taken 3 at a time.

18.7 Random Variables. Discrete and Continuous Distributions

If we roll two dice, we know that the sum X of the two numbers which turn up must be an integer between 2 and 12, but we cannot predict which value of X will occur in the next trial, and we may say that X depends on "chance." Similarly, if we want to draw 5 bolts from a lot of bolts and measure their diameters, we cannot predict how many will be defective, that is, will not meet given requirements; hence $X = \text{number of defectives}$ is again a function which depends on "chance." The lifetime X of a light bulb to be drawn at random from a lot of bulbs also depends on "chance," and so does the content X of a bottle of lemonade filled by a filling machine and selected at random from a given lot.

Roughly speaking, a **random variable** X (also called *stochastic variable* or *variate*) is a function whose values are real numbers and depend on "chance"; more precisely, it is a function X which has the following properties:

1. *The values of X are real numbers.*
2. *To each real number a and each interval I on the real axis there corresponds a well-defined unique probability of the events " X assumes the value a " and " X assumes any value in I ," respectively. These probabilities are in agreement with the axioms in Sec. 18.5.*

Although this definition is quite general and includes many functions, we shall see that in practice the number of important types of random variables and corresponding "probability distributions" is rather small.

If we perform a random experiment and the event corresponding to a number a occurs, then we say that in this trial the random variable X corresponding to that experiment has **assumed** the value a . We also say that we have **observed** the value $X = a$. Instead of "the event corresponding to a number

a ," we say, more briefly, "the event $X = a$." The corresponding probability is denoted by $P(X = a)$. Similarly, the probability of the event

X assumes any value in the interval $a < X < b$

is denoted by $P(a < X < b)$. The probability of the event

$X \leq c$ (X assumes any value smaller than c or equal to c)

is denoted by $P(X \leq c)$, and the probability of the event

$X > c$ (X assumes any value greater than c)

is denoted by $P(X > c)$.

The last two events are mutually exclusive. From Axiom 3, Sec. 18.5, we thus obtain

$$P(X \leq c) + P(X > c) = P(-\infty < X < \infty).$$

From Axiom 2 we see that the right side equals 1, because $-\infty < X < \infty$ is the whole sample space. This yields the important formula

$$(1) \quad P(X > c) = 1 - P(X \leq c) \quad (c \text{ arbitrary, real}).$$

For example, if X is the number that turns up in rolling a fair die, $P(X = 1) = \frac{1}{6}$, $P(X = 2) = \frac{1}{6}$, etc., $P(1 < X < 2) = 0$, $P(1 \leq X \leq 2) = \frac{1}{3}$, $P(0.3 \leq X \leq 3.2) = \frac{1}{2}$, $P(X > 4) = \frac{1}{6}$, $P(X \leq 0.5) = 0$, etc.

In most practical cases the random variables are either *discrete* or *continuous*. These two important notions will now be considered one after the other.

A random variable X and the corresponding distribution are said to be *discrete*, if X has the following properties.

1. The number of values for which X has a probability different from 0 is finite or at most countably infinite.
2. Each finite interval on the real number scale contains at most finitely many of those values. If an interval $a < X \leq b$ does not contain such a value, then $P(a < X \leq b) = 0$.

Let

$$x_1, \quad x_2, \quad x_3, \quad \dots$$

be the values for which X has a positive probability, and let

$$p_1, \quad p_2, \quad p_3, \quad \dots$$

be the corresponding probabilities. Then $P(X = x_1) = p_1$, etc., and the set x_1, x_2, \dots is called the *sample space of X* . We now introduce the function

$$(2) \quad f(x) = \begin{cases} p_j & \text{when } x = x_j, \\ 0 & \text{otherwise.} \end{cases} \quad (j = 1, 2, \dots)$$

$f(x)$ is called the *probability-function of X* .

Since $P(S) = 1$ (cf. Axiom 2 in Sec. 18.5), we must have

$$(3) \quad \sum_{j=1}^{\infty} f(x_j) = 1.$$

If we know the probability function of a discrete random variable X , then we may readily compute the probability $P(a < X \leq b)$ corresponding to any interval $a < X \leq b$. In fact,

$$(4) \quad P(a < X \leq b) = \sum_{a < x_j \leq b} f(x_j) = \sum_{a < x_j \leq b} p_j.$$

This is the sum of all the probabilities $f(x_j) = p_j$, for which the x_j lie in that interval. For a closed, open, or infinite interval the situation is quite similar. We express this by saying that the probability function $f(x)$ determines the **probability distribution** or, briefly, the **distribution** of the random variable X in a unique fashion.

If X is any random variable, not necessarily discrete, then for any real number x there exists the probability $P(X \leq x)$ of the event

$X \leq x$ (X assumes any value smaller than x or equal to x).

Clearly, $P(X \leq x)$ depends on the choice of x ; it is a function of x , which is called the **distribution function**² of X and is denoted by $F(x)$. Thus

$$(5) \quad F(x) = P(X \leq x).$$

Since for any a and $b > a$ we have

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a),$$

it follows that

$$(6) \quad P(a < X \leq b) = F(b) - F(a).$$

This shows that the distribution function determines the distribution of X uniquely and that it can be used for computing probabilities of events.

Suppose that X is a discrete variable. Then we may represent the distribution function $F(x)$ in terms of the probability function $f(x)$. In fact, by inserting (4) (with $a = -\infty$ and $b = x$) we have

$$(7) \quad F(x) = \sum_{x_j \leq x} f(x_j)$$

where the right-hand side is the sum of all those $f(x_j)$ for which $x_j \leq x$. Simple examples are shown in Figs. 389 and 390. In both figures $f(x)$ is represented by a bar chart. In Fig. 389 we have $f(x) = \frac{1}{6}$ when $x = 1, 2, \dots, 6$ and 0 otherwise. In Fig. 390 the function $f(x)$ has the values

x	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

² We note that some authors call $F(x)$ the *cumulative distribution function*, in particular those who use the term *distribution function* for the probability function $f(x)$.

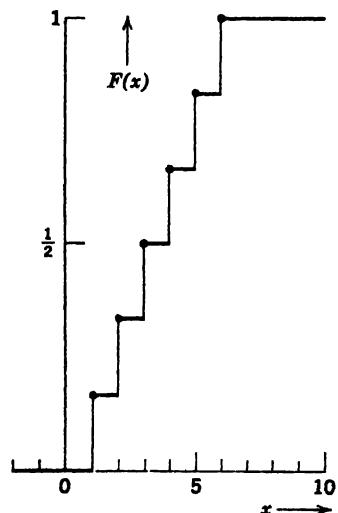
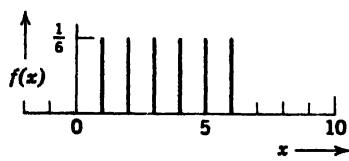


Fig. 389. Probability function $f(x)$ and distribution function $F(x)$ of the random variable X = number obtained in rolling a fair die once.

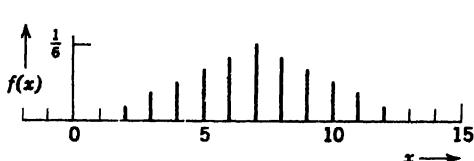


Fig. 390. Probability function $f(x)$ and distribution function $F(x)$ of the random variable X = sum of the two numbers obtained in rolling two fair dice once.

because there are $6 \cdot 6 = 36$ equally likely outcomes, each of which thus has probability $\frac{1}{36}$; the event $X = 2$ contains just one outcome, namely, $(1, 1)$ (where the first number refers to the first die and the second number to the other die); the event $X = 3$ contains the two outcomes $(1, 2)$ and $(2, 1)$; $X = 4$ contains the three outcomes $(1, 3)$, $(2, 2)$, and $(3, 1)$, etc.

The distribution functions $F(x)$ in Figs. 389 and 390 have jumps at those x for which X has a positive probability, the dots indicating the values of $F(x)$ at the x 's corresponding to the jumps. Between two such values of x they are constant. They are so-called *step functions* or *piecewise constant functions*. This is typical for discrete distributions.

More precisely, the values x_1, x_2, x_3, \dots for which a discrete random variable X has positive probability (and only these values) are called the **possible values** of X . In each interval that contains no possible value the distribution function $F(x)$ is constant. Hence $F(x)$ is a step function which has an upward jump of magnitude $p_j = P(X = x_j)$ at $x = x_j$, and is constant between two subsequent possible values.

We shall now define and consider continuous random variables. A random variable X and the corresponding distribution are said to be **continuous** if the

corresponding distribution function $F(x) = P(X \leq x)$ can be represented by an integral in the form

$$(8) \quad F(x) = \int_{-\infty}^x f(v) dv$$

where the integrand is continuous, possibly except for at most finitely many values of v .

Of course, $F(x)$ is continuous for every x . We have used v to denote the variable of integration because x is used to denote the upper limit of integration. The integrand f is called the *probability density* or, briefly, the *density* of the distribution under consideration. Differentiating (8) we see that

$$F'(x) = f(x)$$

for every x at which $f(x)$ is continuous. In this sense *the density is the derivative of the distribution function*.

From (8) and Axiom 2, Sec. 18.5, we also have

$$(9) \quad \int_{-\infty}^{\infty} f(v) dv = 1.$$

Furthermore, from (6) and (8) we obtain

$$(10) \quad P(a < X \leq b) = F(b) - F(a) = \int_a^b f(v) dv.$$

This means that the probability of the event $a < X \leq b$ equals the area under the curve of the density $f(x)$ between $x = a$ and $x = b$, as shown in Fig. 391.

Obviously, for any fixed a and $b (> a)$ the probabilities corresponding to the intervals $a < X \leq b$, $a < X < b$, $a \leq X < b$, and $a \leq X \leq b$ are all the same. This is different from the situation in the case of a discrete distribution.

Since probabilities are non-negative and (10) holds for every interval, we must have $f(x) \geq 0$ for all x .

The distribution function (or equally well the density) of a continuous distribution determines the distribution uniquely. Examples of continuous distributions will be considered later.

We finally mention that there is a strictly quantitative analogy between probability distributions of a single variable and mass distributions of total mass 1 along a straight line. In fact, the analogue of a discrete probability

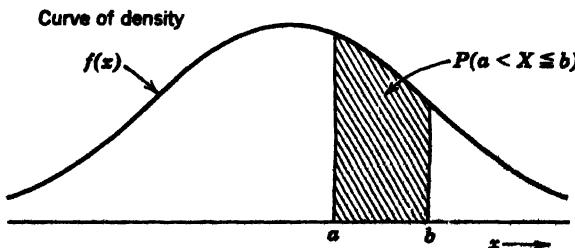


Fig. 391. Example illustrating formula (10).

distribution with probability function (2) is a distribution consisting of masses p_1, p_2, p_3, \dots , located at the points x_1, x_2, x_3, \dots of the x -axis. The mass located at a point x or to the left of it is given by the function

$$F(x) = \sum_{x_i \leq x} f(x_i) = \sum_{x_i \leq x} p_i.$$

This function corresponds to the distribution function $F(x)$ of the probability distribution. The analogue of a continuous probability distribution with density $f(x)$ is a continuous mass distribution of total mass 1 with mass density $f(x)$. The mass located between $-\infty$ and a value x is

$$F(x) = \int_{-\infty}^x f(v) dv,$$

and the mass in an interval $a < x \leq b$ is

$$F(b) - F(a) = \int_a^b f(v) dv.$$

This function $F(x)$ corresponds to the distribution function $F(x)$ of the probability distribution.

Problems

- Find and graph the probability function $f(x)$ of the random variable $X = \text{sum of the 3 numbers obtained in rolling 3 fair dice once.}$
- Show that the random variable $X = \text{number of times a coin is tossed until a head appears}$ has the probability function $f(x) = 2^{-x}$ ($x = 1, 2, \dots$). Show that $f(x)$ satisfies (3).
- Let X be the number of heads obtained in tossing two coins once. Find the probabilities $P(X = 0)$, $P(X = 1)$, $P(X = 2)$, $P(1 < X < 2)$, $P(X \leq 1)$, $P(X \geq 1)$, $P(X > 1)$, and $P(0.5 < X < 10)$.
- A box contains 4 right-handed and 6 left-handed screws. Two screws are drawn at random without replacement. Let X be the number of left-handed screws drawn. Find the probabilities $P(X = 0)$, $P(X = 1)$, etc., as in Prob. 3.
- Find and graph the probability function of the random variable $X = \text{number of heads in tossing 4 coins.}$
- Graph the probability function $f(x) = 0.1x$ when $x = 1, 2, 3, 4$, $f(x) = 0$ otherwise, and the corresponding distribution function.
- Suppose that 2 screws are drawn without replacement from a box containing 10 screws, 4 of which are defective. Graph the probability function $f(x)$ and the distribution function $F(x)$ of the random variable $X = \text{number of defective screws drawn.}$
- What value must the constant k have in order that $f(x) = kx$ when $0 \leq x \leq 2$, $f(x) = 0$ otherwise, is the density of a continuous distribution? Find the corresponding distribution function.
- Consider the density $f(x) = ce^{-ax}$ when $x \geq 0$, $f(x) = 0$ when $x < 0$. Determine c as a function of a . Graph $f(x)$ and the corresponding distribution function $F(x)$ with $a = 0.5$ and $a = 1$.
- Show that $b < c$ implies $P(X \leq b) \leq P(X \leq c)$.

18.8 Mean and Variance of a Distribution

The **mean value** or **mean** of a distribution is denoted by μ and is defined by

$$(1) \quad \begin{aligned} (a) \quad \mu &= \sum_x x f(x), && \text{(discrete distribution),} \\ (b) \quad \mu &= \int_{-\infty}^{\infty} x f(x) dx && \text{(continuous distribution).} \end{aligned}$$

In (1a) the function $f(x)$ is the probability function of the random variable X considered, and we sum over all possible values (cf. Sec. 18.7). In (1b) the function $f(x)$ is the density of X . The mean is also known as the *mathematical expectation of X* and is sometimes denoted by $E(X)$. By definition it is assumed that the series in (1a) converges absolutely and the integral

$$\int_{-\infty}^{\infty} |x| f(x) dx$$

exists. If this does not hold, we say that the distribution does not have a mean; this case will rarely occur in engineering applications.

The mean value corresponds to the center of gravity of the analogous mass distribution. In fact, (1) is the analogue of the well-known formula for that center of gravity.

A distribution is said to be **symmetric** with respect to a number $x = c$ if for every real x ,

$$(2) \quad f(c + x) = f(c - x).$$

The reader may prove (cf. Prob. 1)

Theorem 1. *If a distribution is symmetric with respect to $x = c$ and has a mean μ , then $\mu = c$.*

The **variance** of a distribution is denoted by σ^2 and is defined by

$$(3) \quad \begin{aligned} (a) \quad \sigma^2 &= \sum_x (x - \mu)^2 f(x), && \text{(discrete distribution),} \\ (b) \quad \sigma^2 &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx && \text{(continuous distribution),} \end{aligned}$$

where, by definition, it is assumed that the series in (3a) converges absolutely and the integral in (3b) exists.

In the case of a discrete distribution with $f(x) = 1$ at a point and $f = 0$ otherwise, we have $\sigma^2 = 0$. This case is of no practical interest. In any other case

$$(4)$$

The positive square root of the variance is called the standard deviation and is denoted by σ .

Roughly speaking, the variance is a measure of the spread or dispersion of the values which the variable X considered can assume. From (3) we see that the mechanical analogue of the variance is the moment of inertia with respect to the center of gravity of the corresponding mass distribution.

Example 1. The random variable

$$X = \text{number of heads in a single toss of a coin}$$

has the possible values $X = 0$ and $X = 1$ with probabilities $P(X = 0) = 1/2$ and $P(X = 1) = 1/2$. From (1a) we thus obtain the mean value $\mu = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}$, and (3a) yields

$$\sigma^2 = (0 - \frac{1}{2})^2 \cdot \frac{1}{2} + (1 - \frac{1}{2})^2 \cdot \frac{1}{2} = \frac{1}{4}.$$

Example 2 (Rectangular distribution). The distribution with density

$$f(x) = \frac{1}{b-a} \quad \text{when} \quad a < x < b$$

and $f = 0$ otherwise is called the *uniform* or *rectangular distribution* on the interval $a < x < b$. A particular case is shown in Fig. 392. From Theorem 1 or from (1b) we find that $\mu = (a+b)/2$, and (3b) yields the variance

$$\sigma^2 = \int_a^b \left(x - \frac{a+b}{2} \right)^2 \frac{1}{b-a} dx = \frac{(b-a)^2}{12}.$$

Figure 393 on p. 764 shows special cases illustrating that σ^2 measures the spread.

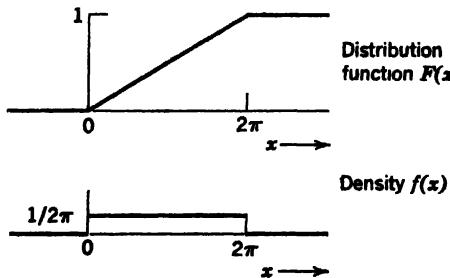


Fig. 392. Density $f(x)$ and distribution function $F(x)$ of the rectangular distribution on the interval $0 < x < 2\pi$.

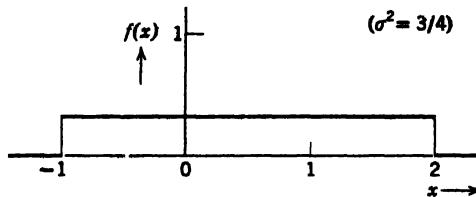
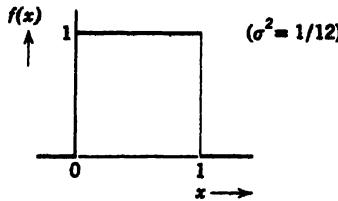
Theorem 2. If a random variable X has mean μ and variance σ^2 , the random variable $X^* = c_1 X + c_2$ ($c_1 \neq 0$) has the mean

$$(5) \quad \mu^* = c_1 \mu + c_2$$

and the variance

$$(6) \quad \sigma^{*2} = c_1^2 \sigma^2$$

Proof. We prove (5) in the continuous case. For corresponding x and $x^* = c_1 x + c_2$ the densities $f(x)$ of X and $f^*(x^*)$ of X^* satisfy the relation



∴ 393. Rectangular distributions having the same mean (1/2) but different varian-

$f^*(x^*) = f(x)/c_1$, because to a small interval of length Δx on the X -axis there corresponds the probability $f(x)\Delta x$ (approximately) and this must equal $f^*(x^*) \Delta x^*$ where $\Delta x^* = c_1 \Delta x$ is the length of the corresponding interval on the X^* -axis. Since $dx^*/dx = c_1$, $dx^* = c_1 dx$, we thus have $f^*(x^*) dx^* = f(x) dx$. Hence

$$\begin{aligned}\mu^* &= \int_{-\infty}^{\infty} x^* f^*(x^*) dx^* = \int_{-\infty}^{\infty} (c_1 x + c_2) f(x) dx \\ &= c_1 \int_{-\infty}^{\infty} x f(x) dx + c_2 \int_{-\infty}^{\infty} f(x) dx.\end{aligned}$$

The last integral equals 1, cf. (9), Sec. 18.7, and (5) is proved. Since

$$x^* - \mu^* = (c_1 x + c_2) - (c_1 \mu + c_2) = c_1 x - c_1 \mu,$$

the definition of the variance yields

$$\sigma^{*2} = \int_{-\infty}^{\infty} (x^* - \mu^*)^2 f^*(x^*) dx^* = \int_{-\infty}^{\infty} (c_1 x - c_1 \mu)^2 f(x) dx = c_1^2 \sigma^2.$$

The proof in the discrete case is similar.

From (5) and (6) we readily obtain

Theorem 3. If a random variable X has mean μ and variance σ^2 , then the corresponding variable $Z = (X - \mu)/\sigma$ has the mean 0 and the variance 1.

Z is called the **standardized variable** corresponding to X .

If X is any random variable and $g(X)$ is any continuous function defined for all real X , then the number

- (a) $E(g(X)) = \sum g(x_i) f(x_i)$ $(X$ discrete)
- (7) (b) $E(g(X)) = \int_{-\infty}^{\infty} g(x) f(x) dx$ $(X$ continuous)

is called the mathematical expectation of $g(X)$. Here f is the probability function or the density, respectively.

Taking $g(X) = X^k$ ($k = 1, 2, \dots$) in (7), we obtain

$$(8) \quad E(X^k) = \sum_j x_j^k f(x_j) \quad \text{and} \quad E(X^k) = \int_{-\infty}^{\infty} x^k f(x) dx,$$

respectively. $E(X^k)$ is called the k th moment of X . Taking $g(X) = (X - \mu)^k$ in (7), we have

$$(9) \quad E([X - \mu]^k) = \sum_j (x_j - \mu)^k f(x_j) \quad \text{and} \quad \int_{-\infty}^{\infty} (x - \mu)^k f(x) dx,$$

respectively. This expression is called the k th central moment of X . The reader may show that

$$(10) \quad E(1) = 1$$

$$(11) \quad \mu = E(X)$$

$$(12) \quad \sigma^2 = E([X - \mu]^2).$$

Problems

1. Prove Theorem 1.
2. Prove (5) and (6) in the discrete case.
3. Derive Theorem 3 from (5) and (6).
4. Show that for any constants a and b ,

$$E(ag(X) + bh(X)) = aE(g(X)) + bE(h(X)).$$

5. Prove (10) and (11).
6. Prove (12).
7. Show that $E(X - \mu) = 0$. What is the analogue of this equation in mechanics?
8. Show that $\sigma^2 = E(X^2) - \mu^2$.
9. Show that $\sigma^2 = E(X[X - 1]) + \mu - \mu^2$.
10. Show that $E([X - \mu]^3) = E(X^3) - 3\mu E(X^2) + 2\mu^3$.
11. (Skewness) The number

$$\gamma = \frac{1}{\sigma^3} E([X - \mu]^3)$$

- is called the *skewness* of X . Justify this term by showing that if X is symmetric with respect to μ and the third central moment exists, this moment is zero.
12. Find the skewness of the distribution with density $f(x) = xe^{-x}$ when $x > 0$ and $f(x) = 0$ otherwise. Graph $f(x)$.
 13. (Moment generating function) The so-called *moment generating function* of a discrete or continuous random variable X is given by the formulas

$$G(t) = E(e^{tX}) = \sum_j e^{tx_j} f(x_j) \quad \text{and} \quad G(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx,$$

respectively. Assuming that differentiation under the summation sign and the integral sign is permissible, show that $\mu = E(X) = G'(0)$ and more generally

$$E(X^k) = G^{(k)}(0)$$

where $G^{(k)}(t)$ is the k th derivative of G with respect to t .

18.9 Binomial, Poisson, and Hypergeometric Distributions

We shall now consider special discrete distributions which are particularly important in statistics. We start with the binomial distribution, which is obtained if we are interested in the number of times an event A occurs in n independent performances of an experiment, assuming that A has probability $P(A) = p$ in a single trial. Then $q = 1 - p$ is the probability that in a single trial the event A does not occur. We assume that the experiment is performed n times and consider the random variable

$$X = \text{number of times } A \text{ occurs.}$$

Then X can assume the values $0, 1, \dots, n$, and we want to determine the corresponding probabilities. For this purpose we consider the event $X = x$ which means that in x of the n trials A occurs and in $n - x$ trials it does not occur. This may look as follows:

$$(1) \quad \underbrace{AA \cdots A}_{x \text{ times}} \quad \underbrace{BB \cdots B}_{n-x \text{ times}}$$

Here $B = \bar{A}$, that is, A does not occur. We assume that the trials are *independent*, that is, do not influence each other. Then, since $P(A) = p$ and $P(B) = q$, we see that (1) has the probability

$$\underbrace{pp \cdots p}_{x \text{ times}} \quad \underbrace{qq \cdots q}_{n-x \text{ times}}$$

Clearly, (1) is merely *one* order of arranging x A 's and $n - x$ B 's, and the probability of $X = x$ thus equals $p^x q^{n-x}$ times the number of different arrangements of x A 's and $n - x$ B 's, as follows from Theorem 1 in Sec. 18.5. We may number the n trials from 1 to n and pick x of these numbers corresponding to those trials in which A happens. Since the order in which we pick the x numbers does not matter, we see from (4a) in Sec. 18.6 that there are $\binom{n}{x}$ different ways in which we can pick those x numbers from the n numbers. Hence the probability $P(X = x)$ of the event $X = x$ equals

$$f(x) = \binom{n}{x} p^x q^{n-x} \quad (x = 0, 1, \dots, n)$$

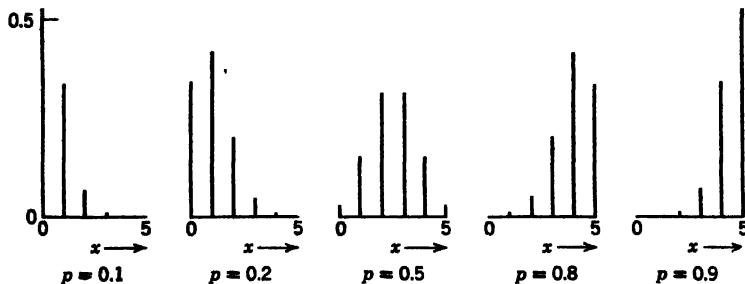


Fig. 394. Probability function (2) of the binomial distribution for $n = 5$ and various values of p .

and $f(x) = 0$ for any other value of x . This is the probability that in n independent trials an event A occurs precisely x times where p is the probability of A in a single trial and $q = 1 - p$. The distribution determined by the probability function (2) is called the **binomial distribution** or *Bernoulli distribution*. The occurrence of A is called *success*, and the nonoccurrence is called *failure*. p is called the *probability of success in a single trial*. Figure 394 shows illustrative examples of $f(x)$.

The binomial distribution has the mean (cf. Prob. 12)

$$(3) \quad \mu = np$$

and the variance (cf. Prob. 13)

$$(4) \quad \sigma^2 = npq.$$

Note that when $p = 0.5$, the distribution is symmetric with respect to μ .

A table of the binomial distribution is included in Sec. 18.21. For more extended tables see Ref. [J6] in Appendix 1.

The distribution with the probability function

$$(5) \quad f(x) = \frac{\mu^x}{x!} e^{-\mu} \quad (x = 0, 1, \dots) \quad$$

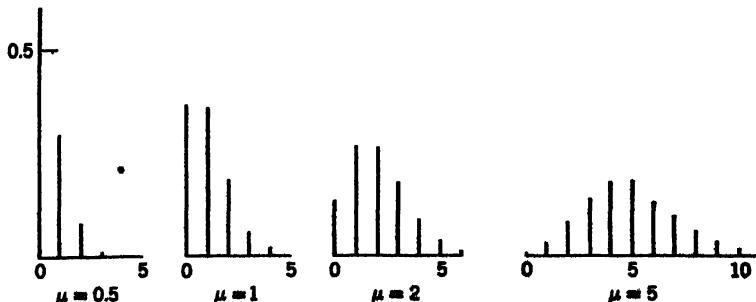


Fig. 395. Probability function (5) of the Poisson distribution for various values of μ .

is called the **Poisson distribution**, named after S. D. Poisson (cf. Sec. 16.5). It can be shown that this distribution may be obtained as a limiting case of the binomial distribution, if we let $p \rightarrow 0$ and $n \rightarrow \infty$ such that the mean $\mu = np$ approaches a finite value. The Poisson distribution has the mean μ and the variance (cf. Prob. 14)

$$(6) \quad \sigma^2 = \mu.$$

This distribution provides probabilities, e.g., of given numbers of cars passing an intersection per unit interval of time, of given numbers of defects per unit length of wire or unit area of paper or textile, and so on.

The binomial distribution is important in **sampling with replacement** (cf. Ex. 3, Sec. 18.5), as follows. Suppose that a box contains N things, for example, screws, M of which are defective. If we want to draw a screw at random, the probability of obtaining a defective screw is

$$p = \frac{M}{N}.$$

Hence in drawing a sample of n screws *with replacement*, the probability that precisely x of the n screws are defective is [cf. (2)]

$$(7) \quad f(x) = \binom{n}{x} \left(\frac{M}{N}\right)^x \left(1 - \frac{M}{N}\right)^{n-x} \quad (x = 0, 1, \dots, n).$$

In **sampling without replacement** that probability is

$$(8) \quad f(x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}} \quad (x = 0, 1, \dots, n).$$

The distribution with the probability function (8) is called the **hypergeometric distribution**.³

To verify our statement we first note that, according to (4a) in Sec. 18.6 there are

- (a) $\binom{N}{n}$ different ways of picking n things from N ,
- (b) $\binom{M}{x}$ different ways of picking x defectives from M ,
- (c) $\binom{N-M}{n-x}$ different ways of picking $n-x$ nondefectives from $N-M$,

and each way in (b) combined with each way in (c) gives the total number of mutually exclusive ways of obtaining x defectives in n drawings without

³ This name comes from the fact that the moment generating function (cf. Prob. 13, Sec. 18.8) of this distribution can be represented in terms of the hypergeometric function.

replacement. Since (a) is the total number of outcomes and we draw at random, each such way has the probability $1/\binom{N}{n}$. From this, (8) follows.

The hypergeometric distribution has the mean (cf. Prob. 16)

$$(9) \quad \mu = n \frac{M}{N}$$

and the variance

$$(10) \quad \sigma^2 = \frac{nM(N - M)(N - n)}{N^2(N - 1)}.$$

Example 1. We want to draw random samples of 2 gaskets from a box containing 10 gaskets, 3 of which are defective. Find the probability function of the random variable

$X = \text{number of defectives in the sample}.$

We have $N = 10$, $M = 3$, $N - M = 7$, $n = 2$. For sampling with replacement, formula (7) yields

$$f(x) = \binom{2}{x} \left(\frac{3}{10}\right)^x \left(\frac{7}{10}\right)^{2-x}, \quad f(0) = 0.49, \quad f(1) = 0.42, \quad f(2) = 0.09.$$

For sampling without replacement we have to use (8), finding

$$f(x) = \binom{3}{x} \binom{7}{2-x} / \binom{10}{2}, \quad f(0) = f(1) = \frac{21}{45} \approx 0.47, \quad f(2) = \frac{3}{45} \approx 0.07.$$

If N , M , and $N - M$ are large compared with n , then it does not matter too much whether we sample with or without replacement, and in this case the hypergeometric distribution may be approximated by the binomial distribution (with $p = M/N$), which is somewhat simpler.

Hence in sampling from an indefinitely large population ("infinite population") we may use the binomial distribution, regardless of whether we sample with or without replacement.

Problems

- Five coins are tossed simultaneously. Find the probability function of the random variable $X = \text{number of heads}$ and compute the probabilities of obtaining no heads, precisely 1 head, at least 1 head, and not more than 4 heads.
- Find formulas for $f(x)$, μ , and σ^2 in the case of the binomial distribution with $p = 0.5$.
- Suppose that for a certain type of aircraft motor the probability of failure of a motor during a 100-hour period is p . Suppose that a 4-motor plane can fly with at least 2 motors, whereas a 2-motor plane can fly with at least 1 motor. Assuming that the failure of one motor is independent of the failure of another and p is small, show that during an intended 100-hour flight the 2-motor plane is about $1/4p$ times more likely to be forced down by motor failure than the 4-motor plane.
- Let $p = 1\%$ be the probability that a certain type of light bulb will fail in a 24-hour test. Find the probability that a sign consisting of 10 such bulbs will burn 24 hours with no bulb failures.

5. A carton contains 20 fuses, 5 of which are defective. Find the probability that, if a sample of 3 fuses is chosen at random from the carton, x will be defective.
6. Suppose that 3% of bolts made by a machine are defective, the defectives occurring at random during production. If the bolts are packaged 50 per box, what is the Poisson approximation of the probability that a given box will contain x defectives?
7. If the probability of hitting a target is 10% and 10 shots are fired independently, what is the probability that the target will be hit at least once?
8. (Multinomial distribution) Suppose a trial can result in precisely one of k mutually exclusive events A_1, \dots, A_k with probabilities p_1, \dots, p_k , respectively, where $p_1 + \dots + p_k = 1$. Suppose that n independent trials are performed. Show that the probability of getting $x_1 A_1$'s, $\dots, x_k A_k$'s is

$$f(x_1, \dots, x_k) = \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k}$$

where $0 \leq x_j \leq n$, $j = 1, \dots, k$, and $x_1 + \dots + x_k = n$. The distribution with this probability function is called the *multinomial distribution*.

9. Suppose that in a production of axle shafts the shafts are classified as *undersize*, *oversize*, and *acceptable* if the diameter measures less than 0.999 in., more than 1.001 in., and between 0.999 and 1.001 in., respectively. Suppose that the production process is such that 3% of the shafts are undersize, 4% are oversize, and 93% are acceptable. If 100 of these shafts are selected at random, what is the probability of getting x_1 undersize, x_2 oversize, and x_3 acceptable shafts?
10. Suppose that in a production of electrical resistors the probability of producing a resistor of resistance $R < 198$ ohms or $R > 201$ ohms is 3% and 5%, respectively. Find the probability that a random sample of 20 resistors contains precisely x_1 resistors with $R < 198$ ohms and precisely x_2 resistors with $R > 201$ ohms.
11. Using the binomial theorem, show that the binomial distribution has the moment generating function (cf. Prob. 13, Sec. 18.8)

$$G(t) = \sum_{x=0}^n e^{tx} \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^n \binom{n}{x} (pe^t)^x q^{n-x} = (pe^t + q)^n.$$

12. Using Prob. 11 and $p + q = 1$, prove (3).
13. Prove (4).
14. Show that the Poisson distribution has the moment generating function

$$G(t) = e^{-\mu} e^{\mu e^t}$$

and prove (6).

15. Using Prob. 14, this section, and Probs. 10 and 11 in Sec. 18.8, show that the Poisson distribution has the skewness $\gamma = 1/\sqrt{\mu}$ and conclude that for large μ the distribution is almost symmetric (cf. Fig. 395).
16. Prove (9).

18.10 Normal Distribution

The continuous distribution having the density

$$(1) \quad f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (\sigma > 0)$$

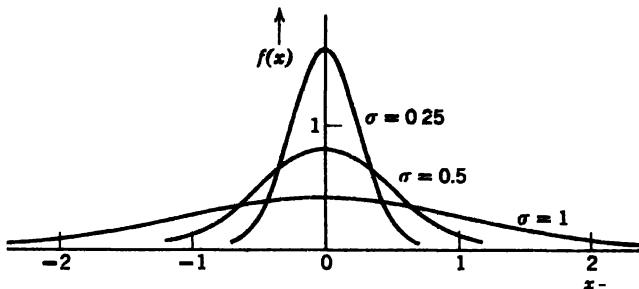


Fig. 396. Density (1) of the normal distribution with $\mu = 0$ for various values of σ .

is called the **normal distribution** or *Gauss distribution*. A random variable having this distribution is said to be **normal** or *normally distributed*. This distribution is very important, because many random variables of practical interest are normal or approximately normal or can be transformed into normal random variables in a relatively simple fashion. Furthermore, certain complicated distributions can be approximated by the normal distribution. This distribution also appears in the mathematical proofs of various statistical tests.

In (1), μ is the mean and σ is the standard deviation of the distribution. The curve of $f(x)$ is called the *bell-shaped curve*. It is symmetric with respect to μ . Figure 396 shows $f(x)$ for $\mu = 0$. For $\mu > 0$ ($\mu < 0$) the curves have the same shape, but are shifted $|\mu|$ units to the right (to the left). The smaller σ^2 is, the higher is the peak at $x = 0$ in Fig. 396 and the steeper are the descents on both sides. This agrees with the meaning of the variance.

From (1) we see that the normal distribution has the distribution function

$$(2) \quad F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{v-\mu}{\sigma}\right)^2} dv.$$

From this and (10) in Sec. 18.7 we obtain

$$(3) \quad P(a < X \leq b) = F(b) - F(a) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}\left(\frac{v-\mu}{\sigma}\right)^2} dv.$$

The integral in (2) cannot be evaluated by elementary methods, but can be represented in terms of the integral (Fig. 397)

$$(4) \quad \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-u^2/2} du,$$

which is the distribution function of the normal distribution with mean 0 and variance 1 and has been tabulated (Table 49 in Sec. 18.21). In fact, if we set $(v - \mu)/\sigma = u$, then $du/dv = 1/\sigma$, $dv = \sigma du$, and we now have to integrate from $-\infty$ to $z = (x - \mu)/\sigma$. From (2) we thus obtain

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{(x-\mu)/\sigma} e^{-u^2/2} \sigma du.$$

σ drops out, and the expression on the right equals (4) where $z = (x - \mu)/\sigma$, that is,

$$(5) \quad F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right).$$

From this important formula and (3) we obtain

$$(6) \quad P(a < X \leq b) = F(b) - F(a) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right).$$

In particular, when $a = \mu - \sigma$ and $b = \mu + \sigma$, the right-hand side equals $\Phi(1) - \Phi(-1)$; to $a = \mu - 2\sigma$ and $b = \mu + 2\sigma$ there corresponds $\Phi(2) - \Phi(-2)$, etc. Using Table 49 in Sec. 18.21 we thus find (Fig. 398)

- (a) $P(\mu - \sigma < X \leq \mu + \sigma) \approx 68\%$,
- (7) (b) $P(\mu - 2\sigma < X \leq \mu + 2\sigma) \approx 95.5\%$,
- (c) $P(\mu - 3\sigma < X \leq \mu + 3\sigma) \approx 99.7\%$.

Hence we may expect that a large number of observed values of a normal random variable X will be distributed as follows:

- (a) About $\frac{2}{3}$ of the values will lie between $\mu - \sigma$ and $\mu + \sigma$.
- (b) About 95% of the values will lie between $\mu - 2\sigma$ and $\mu + 2\sigma$.
- (c) About 99 $\frac{3}{4}$ % of the values will lie between $\mu - 3\sigma$ and $\mu + 3\sigma$.

This may also be expressed as follows.

A value that deviates more than σ from μ will occur about once in 3 trials. A value that deviates more than 2σ or 3σ from μ will occur about once in 20 or 400 trials, respectively. Practically speaking, this means that all the values will lie between $\mu - 3\sigma$ and $\mu + 3\sigma$; these values are called 3-sigma limits.

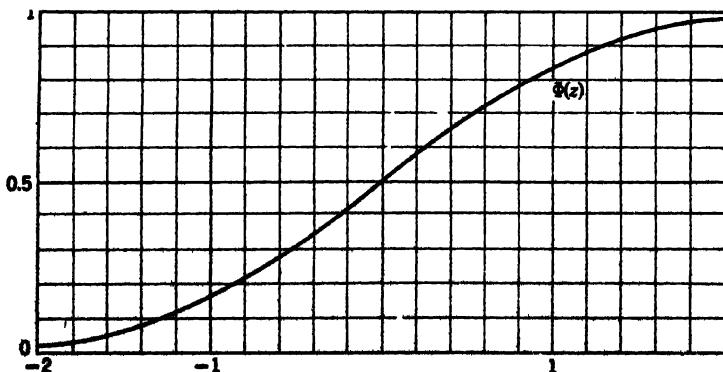


Fig. 397. Distribution function $\Phi(z)$ of the normal distribution with mean 0 and variance 1.

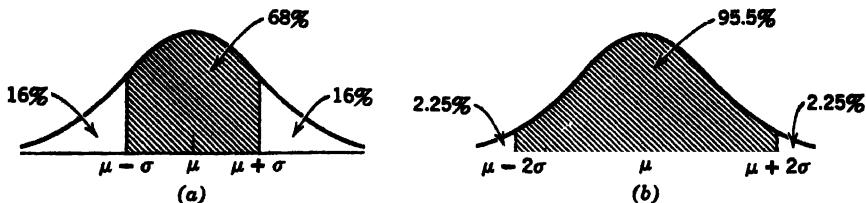


Fig. 398. Illustration of formula (7).

In a similar fashion we obtain

- $$(8) \quad \begin{aligned} (a) \quad P(\mu - 1.96\sigma < X \leq \mu + 1.96\sigma) &= 95\%, \\ (b) \quad P(\mu - 2.58\sigma < X \leq \mu + 2.58\sigma) &= 99\%, \\ (c) \quad P(\mu - 3.29\sigma < X \leq \mu + 3.29\sigma) &= 99.9\%. \end{aligned}$$

The following typical examples should help the reader to understand the practical use of the Tables 49 and 50 in Sec. 18.21.

Example 1. Determine the probabilities

- $$(a) P(X \leq 2.44), \quad (b) P(X \leq -1.16), \quad (c) P(X \geq 1), \quad (d) P(2 \leq X \leq 10)$$

where X is assumed to be normal with mean 0 and variance 1.

Since $\mu = 0$ and $\sigma^2 = 1$, we may obtain the desired values directly from Table 49, finding

- $$\begin{aligned} (a) 0.9927, \quad (b) 0.1230, \quad (c) 1 - P(X \leq 1) &= 1 - 0.8413 = 0.1587 \quad (\text{cf. (6), Sec. 18.5}), \\ (d) \Phi(10) &= 1.0000^\circ (\text{why?}), \quad \Phi(2) = 0.9772, \quad \Phi(10) - \Phi(2) = 0.0228. \end{aligned}$$

Example 2. Determine the probabilities in the previous example, assuming that X is normal with mean 0.8 and variance 4.

From (6) and Table 49 we obtain

$$(a) F(2.44) = \Phi\left(\frac{2.44 - 0.80}{2}\right) = \Phi(0.82) = 0.7939$$

$$(b) F(-1.16) = \Phi(-0.98) = 0.1635$$

$$(c) 1 - P(X \leq 1) = 1 - F(1) = 1 - \Phi(0.1) = 0.4602$$

$$(d) F(10) - F(2) = \Phi(4.6) - \Phi(0.6) = 1 - 0.7257 = 0.2743.$$

Example 3. Let X be normal with mean 0 and variance 1. Determine the constant c such that

- $$\begin{aligned} (a) P(X \geq c) &= 10\%, & (b) P(X \leq c) &= 5\%, \\ (c) P(0 \leq X \leq c) &= 45\%, & (d) P(-c \leq X \leq c) &= 99\%. \end{aligned}$$

From Table 50 in Sec. 18.21 we obtain

$$(a) 1 - P(X \leq c) = 1 - \Phi(c) = 0.1, \quad \Phi(c) = 0.9, \quad c = 1.282$$

$$(b) c = -1.645$$

$$(c) \Phi(c) - \Phi(0) = \Phi(c) - 0.5 = 0.45, \quad \Phi(c) = 0.95, \quad c = 1.645$$

$$(d) c = 2.576.$$

Example 4. Let X be normal with mean -2 and variance 0.25 . Determine c , such that

$$(a) P(X \geq c) = 0.2 \quad (b) P(-c \leq X \leq -1) = 0.5$$

$$(c) P(-2 - c \leq X \leq -2 + c) = 0.9 \quad (d) P(-2 - c \leq X \leq -2 + c) = 99.6\%$$

Using Table 50 in Sec. 18.21, we obtain the following results.

$$(a) 1 - P(X \leq c) = 1 - \Phi\left(\frac{c + 2}{0.5}\right) = 0.2, \Phi(2c + 4) = 0.8, 2c + 4 \Rightarrow 0.842,$$

$$c = -1.579$$

$$(b) \quad \Phi\left(\frac{-1+2}{0.5}\right) - \Phi\left(\frac{-c+2}{0.5}\right) = 0.9772 - \Phi(4 - 2c) = 0.5,$$

$$\Phi(4 - 2c) = 0.4772, 4 - 2c = -0.057, c = 2.03$$

$$(c) \quad \Phi\left(\frac{-2+c+2}{0.5}\right) - \Phi\left(\frac{-2-c+2}{0.5}\right) = \Phi(2c) - \Phi(-2c) = 0.9, \quad 2c = 1.645,$$

c = 0.823

$$(d) \Phi(2c) - \Phi(-2c) = 99.6\%, 2c = 2.878, c = 1.439.$$

Example 5. Suppose that in producing iron plates the plates are required to have a certain thickness, and the machine work is done by a shaper. In any case industrial products will differ slightly from each other because the properties of the material and the behavior of the machines and tools used show slight random variations caused by small disturbances which we cannot predict. We may therefore regard the thickness X [mm] of the plates as a random variable. We assume that for a certain setting the variable X is normal with mean $\mu = 10$ mm and standard deviation $\sigma = 0.02$ mm. We want to determine the percentage of defective plates to be expected, assuming that defective plates are (a) plates thinner than 9.97 mm, (b) plates thicker than 10.05 mm, (c) plates that deviate more than 0.03 mm from 10 mm. (d) How should we choose tolerance limits $10 - c$ and $10 + c$ in order that the expected percentage of defectives be not greater than 5%? (e) How does the percentage of defectives in part (d) change if μ is shifted from 10 mm to 10.01 mm?

Using Table 49 in Sec. 18.21 and (6), we obtain the following solutions.

$$(a) P(X \leq 9.97) = \Phi\left(\frac{9.97 - 10.00}{0.02}\right) = \Phi(-1.5) = 0.0668 \approx 6.7\%$$

$$(b) P(X \geq 10.05) = 1 - P(X \leq 10.05) = 1 - \Phi\left(\frac{10.05 - 10.00}{0.02}\right) \\ = 1 - \Phi(2.5) = 1 - 0.9938 \approx 0.6\%$$

$$(c) P(9.97 \leq X \leq 10.03) = \Phi\left(\frac{10.03 - 10.00}{0.02}\right) - \Phi\left(\frac{9.97 - 10.00}{0.02}\right) \\ = \Phi(1.5) - \Phi(-1.5) = 0.8664. \quad \text{Answer: } 1 - 0.8664 \approx 13\%$$

(d) From (8a) we obtain $c = 1.96\sigma = 0.039$.

Answer: 9.961 mm and 10.039 mm.

$$(e) P(9.961 \leq X \leq 10.039) = \Phi\left(\frac{10.039 - 10.010}{0.02}\right) - \Phi\left(\frac{9.961 - 10.010}{0.02}\right) \\ = \Phi(1.45) - \Phi(-2.45) = 0.9265 - 0.0071 \approx 92\%. \quad \text{Answer: } 8\%.$$

We see that this slight change of the adjustment of the tool bit yields a considerable increase of the percentage of defectives.

Under a linear transformation a normal random variable transforms into a normal random variable, as follows.

Theorem 1. If X is normal with mean μ and variance σ^2 , then $X^* = c_1 X + c_2$ ($c_1 \neq 0$) is normal with mean $\mu^* = c_1 \mu + c_2$ and variance $\sigma^{*2} = c_1^2 \sigma^2$.

Proof. Setting $x = (x^* - c_2)/c_1$, we obtain $(x - \mu)/\sigma = (x^* - \mu^*)/\sigma^*$ where μ^* and σ^* are the expressions in the theorem. X^* has the distribution function

$$\begin{aligned} F(x^*) &= P(X^* \leq x^*) = P(c_1 X + c_2 \leq x^*) = P(X \leq x) \\ &= \Phi\left(\frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x^* - \mu^*}{\sigma^*}\right), \end{aligned}$$

where we made use of (5). This completes the proof.

We shall now see that the normal distribution also yields a useful approximation of the **binomial distribution** which has the probability function

$$(9) \quad f(x) = \binom{n}{x} p^x q^{n-x} \quad (x = 0, 1, \dots, n)$$

(cf. Sec. 18.9), the mean

$$(10) \quad \mu = np,$$

and the variance

$$(11) \quad \sigma^2 = npq.$$

Theorem 2. If n is large, then the binomial distribution can be approximated by the normal distribution with mean (10) and variance (11), that is,

$$f(x) \sim f^*(x) \quad (x = 0, 1, \dots, n),$$

where

$$(12) \quad f^*(x) = \frac{1}{\sqrt{2\pi npq}} e^{-z^2/2}, \quad z = \frac{x - np}{\sqrt{npq}}$$

is the density of that normal distribution and the symbol \sim (read **asymptotically equal**) means that the ratio of both sides approaches 1 as n approaches ∞ . Furthermore,

$$\begin{aligned} (13) \quad P(a \leq X \leq b) &= \sum_{x=a}^b \binom{n}{x} p^x q^{n-x} \sim \Phi(\beta) - \Phi(\alpha), \\ \alpha &= \frac{a - np - 0.5}{\sqrt{npq}}, \quad \beta = \frac{b - np + 0.5}{\sqrt{npq}} \end{aligned}$$

This is the so-called **limit theorem of De Moivre and Laplace**. A proof can be found in Ref. [13], cf. Appendix 1. The proof shows that the term 0.5 in α and β is a correction caused by the change from a discrete to a continuous distribution.

The practical importance of such an approximation is obvious, because for large n the binomial coefficients and powers in (9) become very inconvenient.

Problems

- Let X be normal with mean 100 and variance 36. Find $P(X > 110)$, $P(X < 105)$, and $P(90 < X < 110)$.
- Let X be normal with mean 10 and variance 4. Determine c such that $P(X < c) = 5\%$, $P(X > c) = 1\%$, and $P(10 - c < X < 10 + c) = 50\%$.
- A manufacturer knows from experience that the resistance of resistors he produces is normal with mean $\mu = 100$ ohms and standard deviation $\sigma = 2$ ohms. What percentage of resistors will have resistance between 98 ohms and 102 ohms? Between 95 ohms and 105 ohms?
- In Prob. 3, find c such that 95% of the production lies between $100 - c$ and $100 + c$. Find c^* such that 90% of the production lies between $100 - c^*$ and $100 + c^*$.
- A manufacturer produces airmail envelopes whose weight is normal with mean $\mu = 1.95$ grams and standard deviation $\sigma = 0.05$ grams. The envelopes are sold in lots of 1000. How many envelopes in a lot will be heavier than 2 grams?
- In Example 5, part (c), what value should σ have in order that the percentage of defective plates reduces to 1%?
- Suppose that in a production of fuses the percentage of defectives is 1% and the fuses are sold in lots of 1000 items. What is the probability that such a lot does not contain more than 1% defectives?
- Verify that for $n = 8$, $p = 0.5$, Theorem 2 yields the following approximate values of $f(x)$. Graph these values and the exact values.

x	0	1	2	3	4	5	6	7	8
Approximation	0.005	0.030	0.104	0.220	0.282	0.220	0.104	0.030	0.005
Exact	0.004	0.031	0.109	0.219	0.273	0.219	0.109	0.031	0.004

- What is the probability of obtaining at least 2048 heads if a coin is tossed 4040 times and heads and tails are equally likely? (Cf. Table 35, Sec. 18.5.)

10. Show that

$$(14) \quad \Phi(\infty) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-u^2/2} du = 1.$$

Hint: consider $\Phi^2(\infty)$ and introduce polar coordinates in that double integral.

- Show that the curve of (1) has two points of inflection corresponding to $x = \mu \pm \sigma$.
- Using (14) and integration by parts, show that σ in (1) is the standard deviation of the normal distribution.
- (Bernoulli's law of large numbers). In a random experiment let an event A have probability p ($0 < p < 1$), and let X be the number of times A happens in n independent trials. Show that for any given $\epsilon > 0$,

$$P\left|\left|\frac{X}{n} - p\right| < \epsilon\right| \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

- If X is normal with mean μ and variance σ^2 , what distribution does $-X$ have?

18.11 Distributions of Several Random Variables

If in a random experiment we observe a single quantity, we have to associate with that experiment a single random variable, call it X . From Sec. 18.7 we remember that the corresponding distribution function $F(x) = P(X \leq x)$ determines the distribution completely, because for each event $a < X \leq b$ we have

$$P(a < X \leq b) = F(b) - F(a).$$

If in a random experiment we observe two quantities, we have to associate with the experiment two random variables, say X and Y . For example, X may correspond to the Rockwell hardness and Y to the carbon content of steel, if we measure these two quantities. Each performance of the experiment yields a pair of numbers $X = x$, $Y = y$, briefly (x, y) , which may be plotted as a point in the XY -plane. We may now consider an event $a_1 < X \leq b_1, a_2 < Y \leq b_2$ (Fig. 399), consisting of all the outcomes that lie in this rectangle. If for each such event we know the corresponding probability

$$P(a_1 < X \leq b_1, a_2 < Y \leq b_2),$$

then we say that the **two-dimensional probability distribution** of the random variables X and Y or of the **two-dimensional random variable** (X, Y) is known. The function

$$(1) \quad F(x, y) = P(X \leq x, Y \leq y)$$

is called the **distribution function** of that distribution or of (X, Y) . It determines the distribution uniquely, because (cf. Prob. 1)

$$(2) \quad \begin{aligned} P(a_1 < X \leq b_1, a_2 < Y \leq b_2) \\ = F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2). \end{aligned}$$

The distribution and the variable (X, Y) are said to be **discrete** if (X, Y) has the following properties.

(X, Y) can assume only finitely many or at most countably infinitely many pairs of values (x, y) , the corresponding probabilities being positive. Each

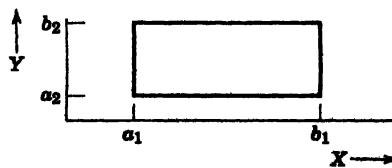


Fig. 399. Notion of a two-dimensional distribution.

bounded domain⁴ contains at most finitely many such pairs, and to any domain containing no such pairs there corresponds the probability 0.

Let (x_i, y_j) be any such pair and let $P(X = x_i, Y = y_j) = p_{ij}$ (where we admit that for certain pairs of subscripts i, j the value p_{ij} may be 0). The function

$$(3) \quad f(x, y) = p_{ij} \quad \text{when } x = x_i, y = y_j, \quad \text{and} \quad f(x, y) = 0 \text{ otherwise}$$

is called the **probability function** of (X, Y) ; here $i = 1, 2, \dots$ and $j = 1, 2, \dots$ independently. The analogue of (7), Sec. 18.7, is

$$(4) \quad F(x, y) = \sum_{x_i \leq x} \sum_{y_j < y} f(x_i, y_j).$$

and instead of (3), Sec. 18.7, we now have

$$(5) \quad \sum_i \sum_j f(x_i, y_j) = 1.$$

For example, if we toss a dime and a nickel once and

$X = \text{number of heads the dime turns up},$

$Y = \text{number of heads the nickel turns up},$

then X and Y can have the values 0 or 1, and

$$f(0, 0) = f(1, 0) = f(0, 1) = f(1, 1) = \frac{1}{4}, \quad f(x, y) = 0 \text{ otherwise.}$$

(X, Y) and its distribution are said to be **continuous** if the corresponding distribution function $F(x, y)$ can be represented by a double integral

$$(6) \quad F(x, y) = \int_{-\infty}^y \int_{-\infty}^x f(x^*, y^*) dx^* dy^*$$

where $f(x, y)$ is defined, nonnegative, and bounded in the entire plane, possibly except for finitely many continuously differentiable curves. $f(x, y)$ is called the **probability density** of the distribution. It follows that

$$(7) \quad P(a_1 < X \leq b_1, a_2 < Y \leq b_2) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y) dx dy.$$

For example (cf. Fig. 400)

$$(8) \quad f(x, y) = 1/k \text{ when } (x, y) \text{ is in } R, \quad f(x, y) = 0 \text{ otherwise}$$

defines the so-called **uniform distribution in the rectangle R** ; here $k = (\beta_1 - \alpha_1)(\beta_2 - \alpha_2)$ is the area of R .

In the case of a **discrete** random variable (X, Y) with probability function $f(x, y)$ we may ask what is the probability $P(X = x, Y \text{ arbitrary})$ of the event that X assume a value x , while Y may assume any value, which we ignore.

⁴ That is, a domain whose points have finite distance from the origin.

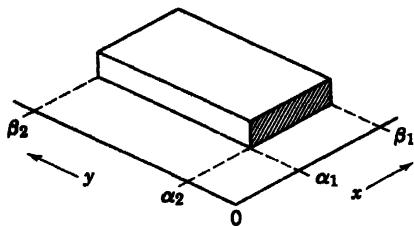


Fig. 400. Probability function (8) of the uniform distribution.

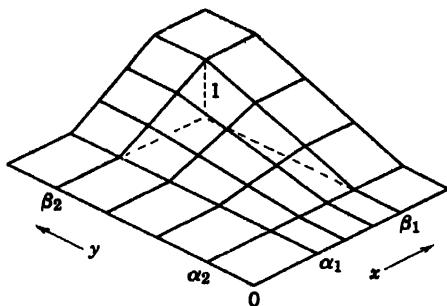


Fig. 401. Distribution function of the uniform distribution defined by (8).

This probability is a function of x , say $f_1(x)$, and we have

$$(9) \quad f_1(x) = P(X = x, Y \text{ arbitrary}) = \sum f(x, y)$$

where we sum all the values of $f(x, y)$ which are not 0 for that x . Clearly, $f_1(x)$ is the probability function of the probability distribution of a single random variable. This distribution is called the **marginal distribution of X with respect to the given two-dimensional distribution**. It has the distribution function

$$(10) \quad F_1(x) = P(X \leq x, Y \text{ arbitrary}) = \sum_{x^* \leq x} f_1(x^*).$$

Similarly, the probability function

$$(11) \quad f_2(y) = P(X \text{ arbitrary}, Y = y) = \sum_x f(x, y)$$

determines the so-called **marginal distribution of Y with respect to the given two-dimensional distribution**. In (10) we sum all the values of $f(x, y)$ which are not 0 for the corresponding y . The distribution function of this distribution is

$$(12) \quad F_2(y) = P(X \text{ arbitrary}, Y \leq y) = \sum_{y^* \leq y} f_2(y^*).$$

Obviously, both marginal distributions of a discrete random variable (X, Y) are discrete.

Table 37 shows an illustrative example. X is the number of queens and Y is the number of kings or aces obtained in drawing 3 cards with replacement from a bridge deck. Since the 52 cards of the deck contain 4 queens, 4 kings, and 4 aces, the probability of obtaining a queen if we draw a single card is $\frac{1}{13} = \frac{1}{3}$, and a king or ace is obtained with probability $\frac{8}{13} = \frac{8}{3}$. Hence to this random experiment there corresponds the probability function

$$f(x, y) = \frac{3!}{x! y! (3-x-y)!} \left(\frac{1}{13}\right)^x \left(\frac{2}{13}\right)^y \left(\frac{10}{13}\right)^{3-x-y} \quad (x + y \leq 3)$$

and $f(x, y) = 0$ otherwise. Table 37 shows the values of $f(x, y)$, $f_1(x)$, and $f_2(y)$.

Table 37. Values of the Probability Functions $f(x, y)$, $f_1(x)$, and $f_2(y)$ in Drawing 3 Cards with Replacement from a Bridge Deck where X is the Number of Queens Drawn and Y is the Number of Kings or Aces Drawn.

$x \ y$	0	1	2	3	$f_1(x)$
0	$\frac{1}{2197}$	$\frac{600}{2197}$	$\frac{120}{2197}$	$\frac{8}{2197}$	$\frac{1728}{2197}$
1	$\frac{300}{2197}$	$\frac{120}{2197}$	$\frac{12}{2197}$	0	$\frac{432}{2197}$
2	$\frac{30}{2197}$	$\frac{6}{2197}$	0	0	$\frac{36}{2197}$
3	$\frac{1}{2197}$	0	0	0	$\frac{1}{2197}$
$f_2(y)$	$\frac{1331}{2197}$	$\frac{726}{2197}$	$\frac{132}{2197}$	$\frac{8}{2197}$	

Similarly, in the case of a *continuous* random variable (X, Y) with density $f(x, y)$ we may consider the event

$$(X \leq x, Y \text{ arbitrary}) \quad \text{or} \quad (X \leq x, -\infty < Y < \infty).$$

Obviously, this event has the probability

$$F_1(x) = P(X \leq x, -\infty < Y < \infty) = \int_{-\infty}^x \left(\int_{-\infty}^{\infty} f(x^*, y) dy \right) dx^*.$$

Setting

$$(13) \quad f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

we may write

$$(14) \quad F_1(x) = \int_{-\infty}^x f_1(x^*) dx^*.$$

$f_1(x)$ is called the *density* and $F_1(x)$ the *distribution function of the marginal distribution of X with respect to the given continuous distribution*. The function

$$(15) \quad f_2(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

is called the *density* and

$$(16) \quad F_2(y) = \int_{-\infty}^y f_2(y^*) dy^* = \int_{-\infty}^y \int_{-\infty}^{\infty} f(x, y^*) dx dy^*$$

is called the *distribution function of the marginal distribution of Y with respect to the given two-dimensional distribution*. We see that both marginal distributions of a continuous distribution are continuous.

The two random variables X and Y of a two-dimensional (X, Y)-distribution with distribution function $F(x, y)$ are said to be *independent* if

$$(17) \quad F(x, y) = F_1(x)F_2(y)$$

holds for all (x, y) . Otherwise these variables are said to be *dependent*.

Suppose that X and Y are either both discrete or both continuous. Then X and Y are independent if and only if the corresponding probability functions or densities $f_1(x)$ and $f_2(y)$ satisfy

$$(18) \quad f(x, y) = f_1(x)f_2(y)$$

for all (x, y) ; cf. Prob. 2. For example, the variables in Table 37 are dependent. The variables $X = \text{number of heads on a dime}$, $Y = \text{number of heads on a nickel}$ in tossing a dime and a nickel once may assume the values 0 or 1 and are independent.

The notions of independence and dependence may be extended to the n random variables of an n -dimensional (X_1, \dots, X_n) -distribution with distribution function

$$F(x_1, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n).$$

These variables are said to be **independent** if

$$(19) \quad F(x_1, \dots, x_n) = F_1(x_1)F_2(x_2) \cdots F_n(x_n)$$

holds for all (x_1, \dots, x_n) . Otherwise these variables are said to be **dependent**. Here $F_j(x_j)$ is the distribution function of the marginal distribution of X_j , that is,

$$F_j(x_j) = P(X_j \leq x_j, X_k \text{ [} k \neq j \text{] arbitrary}).$$

Let (X, Y) be a random variable with probability function or density $f(x, y)$ and distribution function $F(x, y)$, and let $g(x, y)$ be any continuous function defined for all (x, y) . Then $Z = g(X, Y)$ is a random variable, too. For example, if we roll two dice and X is the number that the first die turns up whereas Y is the number that the second die turns up, then $Z = X + Y$ is the sum of those two numbers (cf. Fig. 390 in Sec. 18.7).

If (X_1, \dots, X_n) is an n -dimensional random variable and $g(x_1, \dots, x_n)$ is a continuous function defined for all (x_1, \dots, x_n) , then $Z = g(X_1, \dots, X_n)$ is a random variable, too.

In the case of a *discrete* random variable (X, Y) we may obtain the probability function $f(z)$ of $Z = g(X, Y)$ by summing all $f(x, y)$ for which $g(x, y)$ equals the value of z considered; thus

$$(20) \quad f(z) = P(Z = z) = \sum_{g(x,y)=z} \sum f(x, y).$$

Hence the distribution function of Z is

$$(21) \quad F(z) = P(Z \leq z) = \sum_{g(x,y) \leq z} \sum f(x, y),$$

where we sum all values of $f(x, y)$ for which $g(x, y) \leq z$.

In the case of a *continuous* random variable (X, Y) we similarly have

$$(22) \quad F(z) = P(Z \leq z) = \iint_{g(x,y) \leq z} f(x, y) dx dy,$$

where for each z we integrate over the region $g(x, y) \leq z$ in the xy -plane.

The number

$$(23) \quad E(g(X, Y)) = \begin{cases} \sum_x \sum_y g(x, y) f(x, y) & \text{(discrete distribution)} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy & \text{(continuous distribution)} \end{cases}$$

is called the **mathematical expectation** or, briefly, the **expectation** of $g(X, Y)$. Here it is assumed that the double series converges absolutely and the integral

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |g| f dx dy$$

exists. The formula

$$(24) \quad E(ag(X, Y) + bh(X, Y)) = aE(g(X, Y)) + bE(h(X, Y))$$

can be proved in a fashion similar to that in Prob. 4, Sec. 18.8. An important special case is

$$E(X + Y) = E(X) + E(Y).$$

By induction we obtain from this

Theorem 1 (Addition of means). *The mean (expectation) of a sum of random variables equals the sum of the means (expectations), that is,*

$$(25) \quad E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n).$$

In the case of discrete *independent* random variables X and Y we obtain

$$E(XY) = \sum_x \sum_y xy f(x, y) = \sum_x x f_1(x) \sum_y y f_2(y) = E(X)E(Y),$$

and for continuous *independent* random variables X and Y the proof of $E(XY) = E(X)E(Y)$ is similar. Extension to n random variables yields

Theorem 2 (Multiplication of means). *The mean (expectation) of the product of independent random variables equals the product of the means (expectations), that is,*

$$(26) \quad E(X_1 X_2 \cdots X_n) = E(X_1)E(X_2) \cdots E(X_n).$$

The sum $Z = X + Y$ has the variance (cf. Prob. 8, Sec. 18.8)

$$\sigma^2 = E([Z - \mu]^2) = E(Z^2) - [E(Z)]^2.$$

From (24) it follows that

$$E(Z^2) = E(X^2 + 2XY + Y^2) = E(X^2) + 2E(XY) + E(Y^2),$$

and Theorem 1 yields

$$[E(Z)]^2 = [E(X) + E(Y)]^2 = [E(X)]^2 + 2E(X)E(Y) + [E(Y)]^2.$$

If we insert this, we obtain

$$\begin{aligned}\sigma^2 &= E(X^2) - [E(X)]^2 + E(Y^2) - [E(Y)]^2 \\ &\quad + 2[E(XY) - E(X)E(Y)].\end{aligned}$$

From Prob. 8, Sec. 18.8, we see that the expression in the first line on the right is the sum of the variances of X and Y , which we denote by σ_1^2 and σ_2^2 , respectively. The quantity

$$(27) \quad \sigma_{XY} = E(XY) - E(X)E(Y)$$

is called the **covariance** of X and Y . Consequently,

$$(28) \quad \sigma^2 = \sigma_1^2 + \sigma_2^2 + 2\sigma_{XY}.$$

If X and Y are independent, then $E(XY) = E(X)E(Y)$; hence $\sigma_{XY} = 0$, and

$$(29) \quad \sigma^2 = \sigma_1^2 + \sigma_2^2.$$

Extension to more than two variables yields

Theorem 3 (Addition of variances). *The variance of the sum of independent random variables equals the sum of the variances of these variables.*

Problems

1. Prove (2).
2. Prove the statement involving (18).
3. Using Theorem 1, obtain the formula for the mean μ of the binomial distribution.
4. Using Theorem 3, obtain the formula for the variance σ^2 of the binomial distribution.
5. Graph the probability functions and the distribution functions of the marginal distributions in Table 37.
6. A dime and a nickel are tossed simultaneously. This experiment is performed once. Find and graph the probability function $f(x, y)$, the distribution function $F(x, y)$, and the probability functions and distribution functions of the marginal distributions of (X, Y) where $X = \text{number of heads on the dime}$ and $Y = \text{number of heads on the nickel}$.
7. Find the marginal distributions of the distribution in Figs. 400 and 401 on p. 779.

18.12 Random Sampling. Random Numbers

In the past sections we were concerned with representing statistical data in suitable tabular or graphical forms (Sec. 18.2) and with obtaining mathematical models of populations. In the remaining sections we shall now obtain relations between the theory and the observable reality. More specifically, we shall see what conclusions about a population we can draw if a sample from that population is given and how reliable such conclusions are.

Thus we turn from our consideration of probability to a discussion of basic methods in statistics.

So far it was sufficient to know that a sample from a population is a selection taken from a population (for examples, see Sec. 18.1), but now we have to define this notion in a precise fashion. The sample is supposed to yield information about the corresponding population, because it would be too expensive or time-consuming or even impossible to examine or observe the entire population. In order for information of this type to be obtained, the sample must be a **random selection**. By this we mean that each element of the population must have a known probability of being drawn, that is, of being taken into the sample. If and only if this requirement is satisfied (at least approximately), will the methods to be considered yield sensible results. Otherwise the whole variety of methods would be completely useless.

If the sample space is infinite, the sample values will be **independent**, that is, the results of the n performances of a random experiment made for obtaining n sample values will not influence each other. This certainly applies to samples from a normal population. If the sample space is finite, the sample values will still be independent if we sample with replacement, and they will be *practically* independent if we sample without replacement and the size of the sample is small compared with the size of the population (for example, in the case of samples of 5 or 10 values from a population of 1000 values). However, for large samples the dependence in the case of sampling without replacement from a finite population will matter considerably.

It is not so easy to satisfy the requirement that a sample be a random selection, because there are many and subtle factors of various types that can bias results of sampling. For example, if an interested purchaser wants to draw and inspect a sample of 10 items from a lot of 80 items before he decides whether to purchase the lot, how should he select physically those 10 items so that he can be reasonably sure that all possible $\binom{80}{10}$ samples of size 10 would be equally probable?

Methods to solve this problem have been developed, and we shall now describe such a procedure, which is frequently used.

We number the items of that lot from 1 to 80. Then we select 10 items using Table 51 in Sec. 18.21, which contains *random numbers* made up of sets of **random digits**, as follows. We first select a line number from 0 to 99 at random. This can be done by tossing a coin 7 times denoting heads by 1 and tails by 0, thus generating a 7-place binary number which will represent 0, 1, ..., 127 with equal probabilities. We use this number if it is 0, 1, ..., or 99. Otherwise we ignore it and repeat this process. Then we select a column number from 0 to 9 by generating a 4-place binary number in a similar fashion. Suppose that we obtained 0011010 (= 26) and 0111 (= 7), respectively. In line 26 and column 7 of Table 51 we find 44973. We keep the first 2 digits, that is, 44. We move down the column, starting with 44973 and keeping only the first 2 digits. In this way we find

We omit numbers which are greater than 80 or occur for the second time and continue until 10 numbers are obtained. This yields

44 55 53 03 52 61 67 78 39 54.

The 10 items having these numbers represent the desired selection.

For a larger table of random digits see Ref. [J10], Appendix 1,

Problems

- Suppose that in the example explained in the text we would start from line 83 and column 2 of Table 51, Sec. 18.21, and move upward. What items would then be included in a sample of size 10?
- Using Table 51, Sec. 18.21, select a sample of size 20 from a lot of 250 given items.

18.13 Estimation of Parameters

Constants appearing in distribution functions (e.g., p in the binomial distribution, μ and σ in the normal distribution, etc.) are called **parameters**, and it is a frequent practical task to obtain approximate values of such parameters from a given sample. This task is called *estimation of parameters*.

It is natural to regard the mean \bar{x} of a sample as an approximation of the mean μ of the corresponding population. In this way we have the estimate

$$(1) \quad \mu \approx \bar{x} = \frac{1}{n}(x_1 + \cdots + x_n)$$

where n is the sample size.

Similarly, the variance s^2 of a sample may be regarded as an approximation of the variance σ^2 of the population. Thus

$$(2) \quad \sigma^2 \approx s^2 = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^2.$$

In the case of the binomial distribution we have $p = \mu/n$ (cf. (3), Sec. 18.9). From (1) we thus obtain the estimate

$$(3) \quad p \approx \frac{\bar{x}}{n}.$$

We mention that (1) is a very special case of the so-called **method of moments**. In this method the parameters to be estimated are expressed in terms of the moments of the distribution (cf. Sec. 18.8). In the resulting formulas those moments are replaced by the corresponding moments of the sample. This yields the desired estimates. Here the *kth moment of a sample* x_1, \dots, x_n is

$$m_k = \frac{1}{n} \sum_{j=1}^n x_j^k.$$

In the case of the normal distribution we may obtain estimates of μ and σ by a graphical method, using so-called **normal probability paper**. This paper has the property that the graph of the distribution function of the normal distribution is a straight line. Figure 402 illustrates how the probability paper is obtained from the usual graphical paper by an appropriate stretch of the scale of ordinates. That new scale is constructed by using the inverse

$$(4) \quad w = \Psi(y)$$

of the distribution function $y = \Phi(w)$ defined by (4), Sec. 18.10. Since the normal distribution with mean μ and standard deviation σ has the distribution function (cf. (5), Sec. 18.10)

$$y = F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right),$$

we obtain from (4) a representation of a straight line, namely,

$$(5) \quad w = \Psi(y) = \frac{x - \mu}{\sigma}.$$

$x = \mu$ corresponds to $w = 0$. Since $\Phi(w) = \frac{1}{2}$ when $w = 0$, we thus have the following result.

The mean μ is the abscissa of the point of intersection of the straight line (5) and the 50%-line.

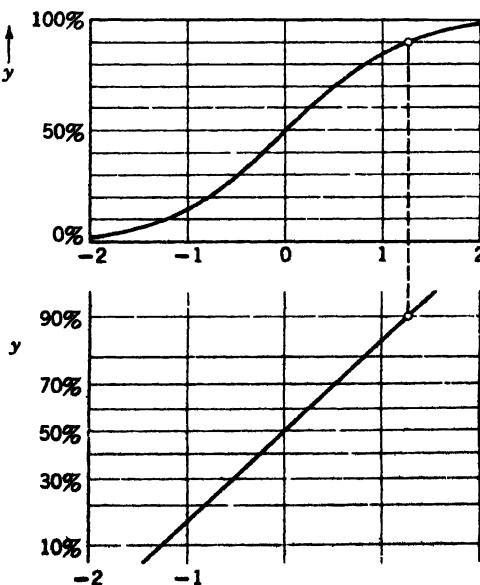


Fig. 402. Curve of the distribution function $y = \Phi(x)$ of the normal distribution with mean 0 and standard deviation 1 and corresponding straight line in probability paper.

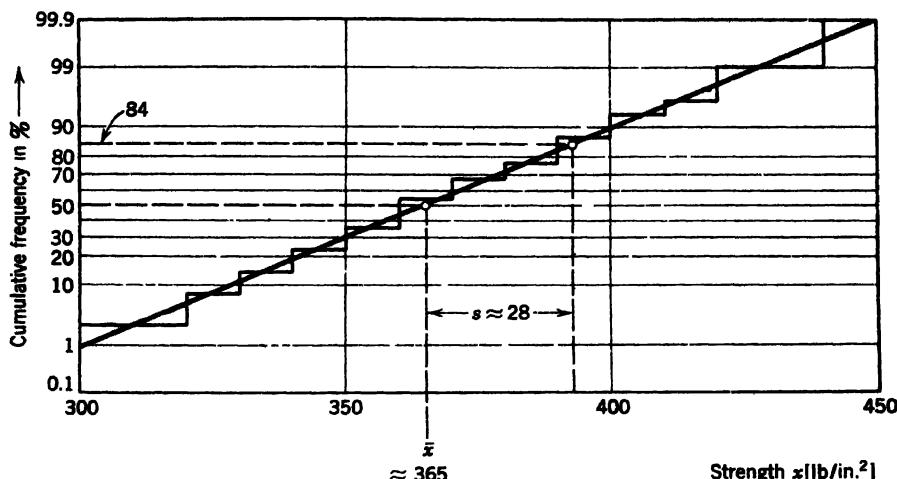


Fig. 403. Distribution function of the sample in Table 28 graphed on probability paper.

Furthermore, $\Phi(1) = 0.84$ (cf. Table 49, Sec. 18.21). Consequently,

$$\Phi\left(\frac{\mu - x}{\sigma}\right) = 0.84 \quad \text{when} \quad \frac{x - \mu}{\sigma} = 1, \quad \text{that is,} \quad x = \mu + \sigma.$$

Hence the straight line (5) intersects the 84%-line approximately at the point with the abscissa $x = \mu + \sigma$.

To determine approximate values (estimates) of μ and σ by the use of a given sample from a normal population we graph the distribution function of the sample as a step curve on the probability paper. Then we draw a straight line which approximates the curve as closely as possible. This line intersects the 50%-line at the point with abscissa $x \approx \bar{x}$. This is an approximation for μ . Furthermore the standard deviation s of the sample is approximately equal to the difference of the x -values of the points of intersection with the 50%- and the 84%-lines. This is an approximation for σ . An example is shown in Fig. 403, where $\bar{x} \approx 365$ (364.7 computed in Table 34, Sec. 18.3) and $s \approx 28$ ($\sqrt{720.1} \approx 26.8$ in Table 34).

If the sample values are equally spaced (as in Table 34), work may be saved by plotting merely the values of the cumulative relative frequencies as points on the probability paper and fitting a straight line as best we can "by eye." From Fig. 403 it is obvious that if we would plot each value above the corresponding x -value, we would make a systematic error because the line would lie higher than that in Fig. 403. To compensate for this effect, we plot each value above the corresponding x -value augmented by $\frac{1}{2}$ the distance between consecutive x -values. (In Table 34 this distance is $\frac{10}{2} = 5$.) This yields Fig. 404 on the next page.

Similarly, if the sample is grouped, we graph the cumulative relative frequencies as points vertically above the right endpoints of the class intervals (not the class midpoints). Then we fit a straight line to these points "by eye."

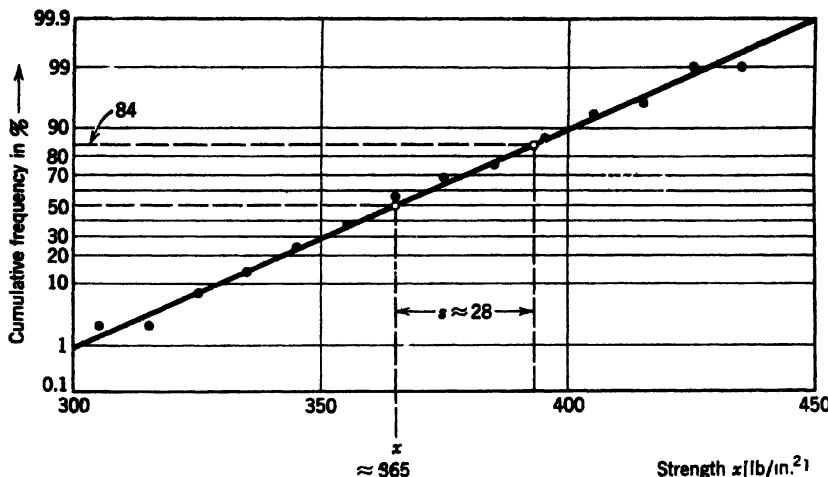


Fig. 404. Cummulative relative frequencies of the sample in Table 28 graphed on probability paper.

Normal probability paper can also be used for obtaining a rough check on whether a sample may have come from a normal population. A test for this problem will be considered in Sec. 18.18.

Another method for obtaining estimates is the so-called **maximum likelihood method** of R. A. Fisher (*Messenger Math.* 41, 1912, 155–160). Let X be a discrete or continuous random variable whose probability function $f(x)$ or density $f(x)$, respectively, depends on a single parameter u . Consider a corresponding sample of n independent values x_1, \dots, x_n . Then in the discrete case the probability that a sample of size n consists precisely of those n values is

$$(6) \quad l = f(x_1)f(x_2) \cdots f(x_n).$$

In the continuous case the probability that the sample consists of values in the small intervals $x_i \leq x \leq x_i + \Delta x$ ($i = 1, 2, \dots, n$) is

$$(7) \quad f(x_1) \Delta x f(x_2) \Delta x \cdots f(x_n) \Delta x = l(\Delta x)^n.$$

Since $f(x_i)$ depends on u , the function l depends on x_1, \dots, x_n and u . We imagine x_1, \dots, x_n to be given and fixed. Then l is a function of u , which is called the **likelihood function**. The basic idea of the maximum likelihood method is very simple, as follows. We choose that approximation for the unknown value of u for which l is as large as possible. If l is a differentiable function of u , a necessary condition for l to have a maximum (not on the boundary) is

$$(8) \quad \frac{\partial l}{\partial u} = 0.$$

(We write a partial derivative, because l depends also on x_1, \dots, x_n .) A solution of (8) depending on x_1, \dots, x_n is called a *maximum likelihood estimate* for u . We may replace (8) by

$$(9) \quad \frac{\partial \ln l}{\partial u} = 0,$$

because $f \geq 0$, a maximum of f is in general positive, and $\ln l$ is a monotone increasing function of l .

If the distribution of X involves r parameters u_1, \dots, u_r , then instead of (8) we have the r conditions

$$(10) \quad \frac{\partial l}{\partial u_1} = 0, \dots, \frac{\partial l}{\partial u_r} = 0,$$

and instead of (9) we now have

$$(11) \quad \frac{\partial \ln l}{\partial u_1} = 0, \dots, \frac{\partial \ln l}{\partial u_r} = 0.$$

Example 1 (Normal distribution). Find maximum likelihood estimates for μ and σ in the case of the normal distribution. From (1), Sec. 18.10, and (6) we obtain

$$l = \left(\frac{1}{\sqrt{2\pi}} \right)^n \left(\frac{1}{\sigma} \right)^n e^{-h} \quad \text{where} \quad h = \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2.$$

Taking logarithms, we have

$$\ln l = -n \ln \sqrt{2\pi} - n \ln \sigma - h.$$

The first equation in (11) is $\partial \ln l / \partial \mu = 0$, at length

$$\frac{\partial \ln l}{\partial \mu} = -\frac{\partial h}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0, \quad \text{hence} \quad \sum_{i=1}^n x_i - n\mu = 0.$$

The solution is the desired estimate $\bar{\mu}$ for μ ; we find

$$\bar{\mu} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}.$$

The second equation in (11) is $\partial \ln l / \partial \sigma = 0$, at length

$$\frac{\partial \ln l}{\partial \sigma} = -\frac{n}{\sigma} - \frac{\partial h}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (x_i - \mu)^2 = 0.$$

Replacing μ by $\bar{\mu}$ and solving for σ^2 , we obtain the estimate

$$\tilde{s}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2,$$

which we shall use in Sec. 18.18. Note that this differs from (2). We cannot discuss criteria for the goodness of estimates, but we want to mention that for small n , formula (2) is preferable.

Problems

Using probability paper, find estimates for μ and σ from the given sample.

1. The sample in Ex. 1, Sec. 18.3.
2. The (ungrouped) sample in Table 29, Sec. 18.2.
3. The grouped sample in Table 30, Sec. 18.2.
4. The sample in Prob. 4, Sec. 18.2.
5. The sample in Prob. 5, Sec. 18.2.

18.14 Confidence Intervals

In the last section we considered approximations for parameters in probability distributions. A numerical value obtained from such a formula by the use of a sample from the population is called a *point estimate* of the corresponding parameter. There is another type of estimate, called *interval estimate*, which we shall now consider, starting with its general motivation.

Whenever we use mathematical approximation formulas, we should try to find out how much the approximate value can at most deviate from the unknown true value. For example, in the case of numerical integration methods there exist "error formulas" from which one can compute the maximum possible error (= difference between the approximate and the true value; cf. Sec. 0.8). Suppose that in a certain case we have obtained 2.47 as an approximate value of a given integral and ± 0.02 as the maximum possible deviation from the unknown exact value. Then we are sure that the values $2.47 - 0.02 = 2.45$ and $2.47 + 0.02 = 2.49$ "include" the unknown exact value, that is, 2.45 is smaller than or equal to that value and 2.49 is larger than or equal to that value.

In the case of an estimation of a parameter u the corresponding problem would be the determination of two numerical quantities that depend on the sample values and include the unknown value of the parameter with certainty. However, we already know that from a sample we cannot draw conclusions about the corresponding population that are 100% certain. So we have to be more modest and have to modify our problem, as follows.

Choose a probability γ close to 1 (for example, $\gamma = 95\%$, 99% , or the like). Then determine two quantities \bar{U}_1 and \bar{U}_2 such that the probability that \bar{U}_1 and \bar{U}_2 include the exact unknown value of the parameter u is equal to γ .

The point is that the request "with certainty" is replaced by the request "with a preassigned probability close to 1."

Those two values should be computed from a given sample x_1, \dots, x_n . The n sample values may be regarded as observed values of n random variables X_1, \dots, X_n . Then \bar{U}_1 and \bar{U}_2 are functions of these random variables and therefore random variables, too. Our above request may thus be written

$$\bullet \quad P(\bar{U}_1 \leq u \leq \bar{U}_2) = \gamma.$$

If we know such functions \tilde{U}_1 and \tilde{U}_2 and a sample is given, we may compute a numerical value u_1 of \tilde{U}_1 and a numerical value u_2 of \tilde{U}_2 . The interval with endpoints u_1 and u_2 is called a **confidence interval** or *interval estimate* for the unknown parameter u , and we shall denote it by $\text{CONF } \{u_1 \leq u \leq u_2\}$. The values u_1 and u_2 are called **confidence limits**. The number γ is called the **confidence level**. One chooses $\gamma = 95\%$, 99% , sometimes 99.9% .

Clearly, if we intend to obtain a sample and determine a corresponding confidence interval, then γ is the probability of getting an interval that will include the unknown exact value of the parameter.

For example, if we choose $\gamma = 95\%$, then we may expect that *about* 95% of the samples that we may obtain will yield confidence intervals that do include the value u , while the remaining 5% do not. Hence the statement "the confidence interval includes u " will be correct in *about* 19 out of 20 cases, while in the remaining case it will be false.

Choosing $\gamma = 99\%$ instead of 95%, we may expect the statement to be correct even in *about* 99 out of 100 cases. But we shall see later that the intervals corresponding to $\gamma = 99\%$ are longer than those corresponding to $\gamma = 95\%$. This is the disadvantage of increasing γ .

What value of γ should we choose in a concrete case? This is not a mathematical question, but one that must be answered from the viewpoint of the application, taking into account what risk of making a false statement we can afford.

We shall now consider methods for obtaining confidence intervals for the mean (Tables 38, 39) and the variance (Table 40) of the normal distribution. The corresponding theory will be presented in the last part of this section.

Table 38. Determination of a Confidence Interval for the Mean μ of a Normal Distribution with Known Variance σ^2 .

1st step. Choose a confidence level γ (95%, 99% or the like).

2nd step. Determine the corresponding c :

γ	0.90	0.95	0.99	0.999
c	1.645	1.960	2.576	3.291

3rd step. Compute the mean \bar{x} of the sample x_1, \dots, x_n .

4th step. Compute $k = c\sigma/\sqrt{n}$. The confidence interval for μ is

$$(1) \quad \text{CONF } \{\bar{x} - k \leq \mu \leq \bar{x} + k\}.$$

Example 1. Determine a 95%-confidence interval for the mean of a normal distribution with variance $\sigma^2 = 9$, using a sample of $n = 100$ values with mean $\bar{x} = 5$.

1st step. $\gamma = 0.95$ is required.

2nd step. The corresponding c equals 1.960.

3rd step. $\bar{x} = 5$ is given.

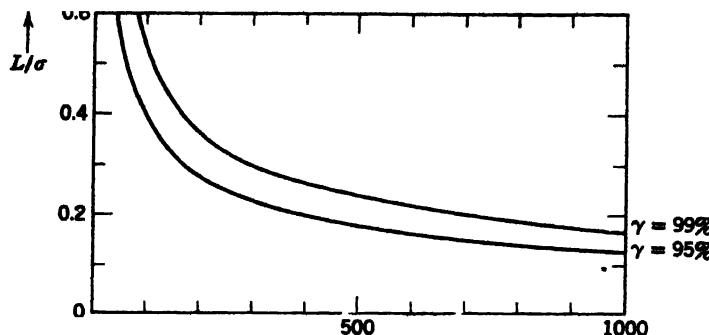


Fig. 405. Length of the confidence interval (1) (measured in multiples of σ) as a function of the sample size n .

4th step. We need $k = 1.960 \cdot 3/\sqrt{100} = 0.588$. Hence $\bar{x} - k = 4.412$, $\bar{x} + k = 5.588$, and the confidence interval is

$$\text{CONF } \{4.412 \leq \mu \leq 5.588\}.$$

Sometimes this is written $\mu = 5 \pm 0.588$, but we shall not use this notation, which may be misleading.

Example 2 (Determination of the sample size necessary for obtaining a confidence interval of prescribed length). How large must n in the last example be, if we want to obtain a 95%-confidence interval of length $L = 0.4$?

The interval (1) has the length $L = 2k = 2c\sigma/\sqrt{n}$. Solving for n , we obtain

$$n = (2c\sigma/L)^2.$$

In the present case the answer is $n = (2 \cdot 1.960 \cdot 3/0.4)^2 \approx 870$.

Figure 405 shows that L decreases as n increases. The shorter the confidence interval should be, the larger the sample size n must be chosen. Figure 406 shows how L increases with increasing γ (and constant n). Clearly $L \rightarrow \infty$ as $\gamma \rightarrow 1$.

Table 39 shows how to determine a confidence interval for the mean μ of a normal distribution with unknown variance σ^2 . The steps are similar to those

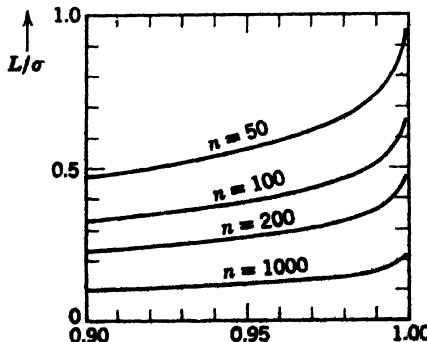


Fig. 406. Length of the confidence interval (1) (measured in multiples of σ) as a function of the confidence level γ for constant n .

Table 39. Determination of a Confidence Interval for the Mean μ of a Normal Distribution with Unknown Variance σ^2 .

1st step. Choose a confidence level γ (95%, 99% or the like).

2nd step. Determine the solution c of the equation

$$(2) \quad F(c) = \frac{1}{2}(1 + \gamma)$$

from the table of the t -distribution with $n - 1$ degrees of freedom (Table 52 in Sec. 18.21; n = sample size).

3rd step. Compute the mean \bar{x} and the variance s^2 of the sample x_1, \dots, x_n .

4th step. Compute $k = sc/\sqrt{n}$. The confidence interval is

$$(3) \quad \text{CONF} \{ \bar{x} - k \leq \mu \leq \bar{x} + k \}.$$

in Table 38. The difference is that k is different from k in Table 38 and c depends on n and must be determined from Table 52 in Sec. 18.21. This table contains values z corresponding to given values of the distribution function

$$(4) \quad F(z) = K_m \int_{-\infty}^z \left(1 + \frac{u^2}{m}\right)^{-(m+1)/2} du$$

of the so-called **t-distribution** of Student (pseudonym for W. S. Gosset) where $K_m = \Gamma(\frac{1}{2}m + \frac{1}{2})/[\sqrt{m\pi}\Gamma(\frac{1}{2}m)]$, and $\Gamma(\alpha)$ is the gamma function (Sec. 17.1). $m (= 1, 2, \dots)$ is a parameter which is called the *number of degrees of freedom* of the distribution.

Example 3. Using the sample in Table 28, Sec. 18.2, determine a 99%-confidence interval for the mean μ of the corresponding population, assuming that the population is normal. (This assumption will be justified in Sec. 18.18.)

1st step. $\gamma = 0.99$ is required.

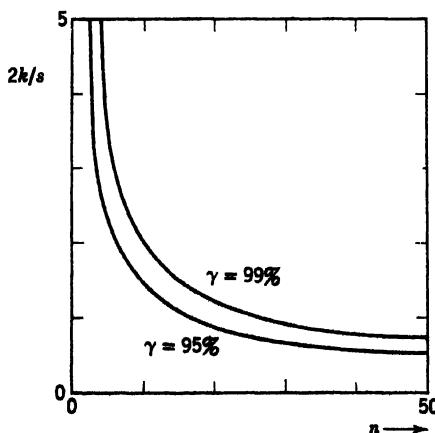


Fig. 407. Length of the confidence interval (3) measured in multiples of s .

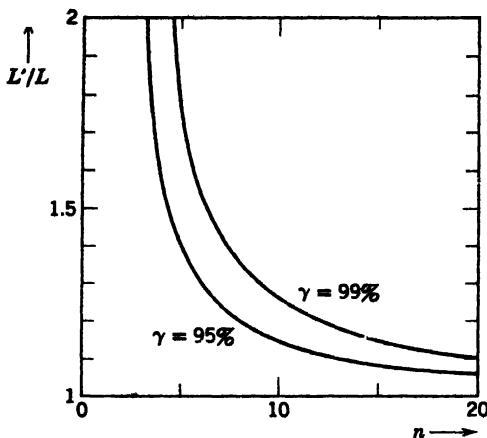


Fig. 408. Ratio of the lengths L' and L of the confidence intervals (3) and (1) as a function of the sample size n for equal s and σ .

2nd step. Since $n = 100$, we obtain $c = 2.63$.

3rd step. Table 34 in Sec. 18.3 shows that $\bar{x} = 364.70$ and $s = \sqrt{720.1} = 26.83$.

4th step. We find $k = 26.83 \cdot 2.63/10 = 7.06$. Hence the confidence interval is

$$\text{CONF } \{357.64 \leq \mu \leq 371.76\}.$$

Table 40. Determination of a Confidence Interval for the Variance σ^2 of the Normal Distribution, Whose Mean Need Not Be Known.

1st step. Choose a confidence level γ (95%, 99% or the like).

2nd step. Determine solutions c_1 and c_2 of the equations

$$(5) \quad F(c_1) = \frac{1}{2}(1 - \gamma), \quad F(c_2) = \frac{1}{2}(1 + \gamma)$$

from the table of the chi-square distribution with $n - 1$ degrees of freedom (Table 53 in Sec. 18.21; n = size of sample).

3rd step. Compute $(n - 1)s^2$ where s^2 is the variance of the sample x_1, \dots, x_n .

4th step. Compute $k_1 = (n - 1)s^2/c_1$ and $k_2 = (n - 1)s^2/c_2$. The confidence interval is

$$(6) \quad \text{CONF } \{k_2 \leq \sigma^2 \leq k_1\}.$$

The steps in Table 40 for obtaining a confidence interval for the variance of a normal distribution are similar to those in Tables 38 and 39, but we now have to determine two numbers c_1 and c_2 . Both numbers are obtained from Table 53 in Sec. 18.21, which contains values z corresponding to given values of the distribution function

$$(7) \quad F(z) = \int_0^z e^{-u/2} u^{(m-2)/2} du \quad \text{when } z \geq 0,$$

$F(z) = 0$ when $z < 0$, of the so-called **χ^2 -distribution** (*chi-square distribution*) where $C_m = 1/[2^{m/2}\Gamma(\frac{1}{2}m)]$ and $m (= 1, 2, \dots)$ is a parameter which is called the *number of degrees of freedom* of the distribution.

Example 4. Using the sample in Table 28, Sec. 18.2, determine a 95%-confidence interval for the variance of the corresponding population.

1st step. $\gamma = 0.95$ is required.

2nd step. Since $n = 100$, we find $c_1 = 73.4$ and $c_2 = 128$.

3rd step. From Table 34 in Sec. 18.3 we see that $99s^2 = 71,291$.

4th step. The confidence interval is

$$\text{CONF } \{556 \leq \sigma^2 \leq 972\}.$$

Other distributions. Confidence intervals for the mean and variance of other distributions may be obtained by using the previous methods and *sufficiently large samples*. Practically speaking, if the sample indicates that the skewness of the unknown distribution is small, one should take samples of size $n = 20$ at least for obtaining confidence intervals for μ and samples of size $n = 50$ at least for obtaining confidence intervals for σ^2 . The reason for this method will be explained at the end of this section.

We shall now consider the theory which justifies our methods for obtaining confidence intervals, using the following simple but very important idea.

So far we have regarded the values x_1, \dots, x_n of a sample as n observed values of a single random variable X . We may equally well regard these n values as single observations of n random variables X_1, \dots, X_n which have the same distribution (the distribution of X) and are independent because the sample values are assumed to be independent.

For deriving (1) in Table 38 we need

Theorem 1 (Sum of independent normal random variables). Suppose that X_1, X_2, \dots, X_n are independent normal random variables with means $\mu_1, \mu_2, \dots, \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$, respectively. Then the random variable

$$X = X_1 + X_2 + \dots + X_n$$

is normal with the mean

$$\mu = \mu_1 + \mu_2 + \dots + \mu_n$$

and the variance

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2.$$

The statements about μ and σ follow directly from Theorems 1 and 3 in Sec. 18.11. The proof that X is normal can be found in Ref. [I3], cf. Appendix 1.

From this theorem, Theorem 1 in Sec. 18.10, and Theorem 3 in Sec. 18.8 we obtain

Theorem 2. If X_1, \dots, X_n are independent normal random variables each of which has mean μ and variance σ^2 , then

$$(8) \quad \bar{X} = \frac{1}{n}(X_1 + \dots + X_n)$$

is normal with the mean μ and the variance σ^2/n , and

$$(9) \quad Z = \sqrt{n} \frac{\bar{X} - \mu}{\sigma}$$

is normal with the mean 0 and the variance 1.

We choose a number γ between 0 and 1 and determine c from Table 50, Sec. 18.21, such that $P(-c \leq Z \leq c) = \gamma$. (When $\gamma = 0.90$, etc., we obtain the values c in Table 38.) The inequality $-c \leq Z \leq c$, at length

$$-c \leq \sqrt{n}(\bar{X} - \mu)/\sigma \leq c$$

[cf. (9)] can be transformed into an inequality for μ , as follows. Multiplication by σ/\sqrt{n} yields $-k \leq \bar{X} - \mu \leq k$ where $k = c\sigma/\sqrt{n}$. Multiplying by -1 and adding \bar{X} , we get

$$(10) \quad \bar{X} + k \geq \mu \geq \bar{X} - k.$$

Thus $P(-c \leq Z \leq c) = \gamma$ is equivalent to $P(\bar{X} - k \leq \mu \leq \bar{X} + k) = \gamma$. Regarding the sample values x_1, \dots, x_n used in Table 38 as observed values of n independent normal random variables X_1, \dots, X_n , we see that the sample mean \bar{x} is an observed value of (8). By inserting this value into (10) we obtain formula (1).

For deriving (3) in Table 39 we need

Theorem 3. Let X_1, \dots, X_n be independent normal random variables with the same mean μ and the same variance σ^2 . Then the random variable

$$(11) \quad T = \sqrt{n} \frac{\bar{X} - \mu}{S}$$

where \bar{X} is given by (8) and

$$(12) \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

has a *t*-distribution [cf. (4)] with $n - 1$ -degrees of freedom.

The proof can be found in Ref. [13], cf. Appendix 1.

We choose a number γ between 0 and 1 and determine a number c from Table 52, Sec. 18.21, with $n - 1$ degrees of freedom such that

$$(13) \quad P(-c \leq T \leq c) = F(c) - F(-c) = \gamma.$$

Since the *t*-distribution is symmetric, we have $F(-c) = 1 - F(c)$, and (13) assumes the form (2). Transforming $-c \leq T \leq c$ in (13) as before, we get

$$(14) \quad \bar{X} - K \leq \mu \leq \bar{X} + K \quad \text{where} \quad K = cS/\sqrt{n},$$

and (13) becomes $P(\bar{X} - K \leq \mu = \bar{X} + K) = \gamma$. By inserting the observed values \bar{x} of \bar{X} and s^2 of S^2 into (14) we obtain (3).

For deriving (6) in Table 40 we need

Theorem 4. Under the assumptions in Theorem 3 the random variable

$$(15) \quad Y = (n - 1) \frac{S^2}{\sigma^2}$$

where S^2 is given by (12), has a chi-square distribution [cf. (7)] with $n - 1$ degrees of freedom.

The proof can be found in Ref. [I3], cf. Appendix 1.

We choose a number γ between 0 and 1 and determine c_1 and c_2 from Table 53, Sec. 18.21, such that [cf. (5)]

$$P(Y \leq c_1) = F(c_1) = \frac{1}{2}(1 - \gamma), \quad P(Y \leq c_2) = F(c_2) = \frac{1}{2}(1 + \gamma).$$

Subtraction yields

$$P(c_1 \leq Y \leq c_2) = P(Y \leq c_2) - P(Y \leq c_1) = \gamma.$$

Transforming $c_1 \leq Y \leq c_2$ with Y given by (15) into an inequality for σ^2 , we obtain

$$\frac{n-1}{c_2} S^2 \leq \sigma^2 \leq \frac{n-1}{c_1} S^2$$

By inserting the observed value s^2 of S^2 we obtain (6).

We know that if X_1, \dots, X_n are independent random variables with the same mean μ and the same variance σ^2 , then their sum $X = X_1 + \dots + X_n$ has the following properties.

(A) X has the mean $n\mu$ and the variance $n\sigma^2$ (cf. Theorems 1 and 3 in Sec. 18.11).

(B) If those variables are normal, then X is normal (cf. Theorem 1).

If those variables are not normal, then (B) fails to hold, but if n is large, then X is approximately normal in the sense of the following basic theorem.

Theorem 5 (Central limit theorem). Let X_1, \dots, X_n, \dots be independent random variables that have the same distribution function and therefore the same mean μ and the same variance σ^2 . Let $Y_n = X_1 + \dots + X_n$. Then the random variable

$$(16) \quad Z_n = \frac{Y_n - n\mu}{\sigma\sqrt{n}}$$

is asymptotically normal with the mean 0 and the variance 1, that is, the distribution function $F_n(x)$ of Z_n satisfies

$$\lim_{n \rightarrow \infty} F_n(x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du.$$

A proof can be found in Ref. [13], cf. Appendix 1. This theorem justifies the application of methods for the normal distribution to other distributions, using large samples.

Problems

- Suppose that 10, 12, 15, 16, 19 is a sample taken from a normal population with variance $\sigma^2 = 6.25$. Find a 95%-confidence interval for the mean μ .
- Find a 99%-confidence interval for the mean μ of a normal population with variance $\sigma^2 = 5.76$ using a sample of size $n = 225$ with sample mean $\bar{x} = 7.26$.
- What sample size must be used in Prob. 2 for obtaining a 99%-confidence interval of length 0.1? For obtaining a 99.9%-confidence interval of length 0.1?
- Graph the length L of the confidence interval (1) with $\gamma = 99.9\%$ as a function of n and compare the result with Fig. 405.
- The confidence interval (1) is a special case of a more general confidence interval which we obtain if we replace $P(-c \leq Z \leq c) = \gamma$ by

$$P(-c_1 \leq Z \leq c_2) = \gamma,$$

where c_1 is suitably chosen and c_2 is obtained from Table 50 in Sec. 18.21. Show that the more general confidence interval is

$$\text{CONF} \left\{ \bar{x} - c_2 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + c_1 \frac{\sigma}{\sqrt{n}} \right\}$$

- For fixed γ , σ , and n the length L of the confidence interval in Prob. 5 depends on c_1 and c_2 . Show that L is minimum if and only if $c_1 = c_2$ [in which case that confidence interval is identical with (1)].
- Using Theorem 1, find out in what fraction of random matches of holes and bolts the bolts would fit, assuming that the diameter D_1 of the holes is normal with mean $\mu_1 = 1$ in. and standard deviation $\sigma_1 = 0.01$ in. and the diameter D_2 of the bolts is normal with mean $\mu_2 = 0.99$ in. and standard deviation $\sigma_2 = 0.01$ in.
- Suppose that the content X of boxes of powder packaged by a machine is normal with mean 100 grams and standard deviation 1 gram and the weight Y of the empty boxes is normal with mean 5 grams and standard deviation 0.5 gram. Find the percentage of packaged boxes whose weight lies between 104 grams and 106 grams.

Assuming that the populations from which the following samples are taken are normal, determine a 95%-confidence interval for the mean μ of the population.

- Nitrogen content [%] of steel

0.74 0.75 0.73 0.75 0.74 0.72

- Density [grams per cm³] of coke

1.40 1.45 1.39 1.44 1.38

- Weight [grams] of duplicating paper

4.3 4.5 4.2 4.3 4.3 4.7 4.4 4.2 4.3 4.5

- Length [mm] of matches

44.4 44.2 44.3 44.2 44.6 44.4 44.7 44.3 44.5 43.9

- What sample size should we use in Prob. 11 if we want to obtain a confidence interval of length 0.1 gram?

Assuming that the populations from which the following samples are taken are normal, determine a 99 %-confidence interval for the variance σ^2 of the population.

14. The sample in Prob. 9

15. Rockwell hardness of tool bits

64.9 64.1 63.8 64.0

16. A sample of size $n = 128$ (tensile strength of sheet steel, measured in kg/mm²) with variance $s^2 = 1.921$

17. Weight [grams] of Theragran tablets

1.19 1.23 1.18 1.21 1.27 1.17 1.15 1.14 1.19 1.20

18.15 Testing of Hypotheses, Decisions

A statistical **hypothesis** is an assumption about the distribution of a random variable, for example, that a certain distribution has mean 20.3, etc. A statistical test of a hypothesis is a procedure in which a sample is used to find out whether we can **accept** the hypothesis, that is, believe that it is true, or whether we should "**reject**" it, that is, believe that it is false.

Statistical tests are of great practical importance. A typical situation is that one has a choice between two possibilities, and the decision which way to take is made on the basis of the result of a test. So the whole subject matter may actually be regarded as a *theory of decisions*.

For example, if we want to use a certain lathe for producing bolts whose diameter should lie between given tolerance limits and we allow at most 2% defective bolts, we may take a sample of 100 bolts produced on that lathe and use it for testing the hypothesis $\sigma^2 = \sigma_0^2$ that the variance σ^2 of the corresponding population has a certain value σ_0^2 which we choose such that we can expect to obtain not more than 2% defectives. If the test leads to the acceptance of the hypothesis, we shall use the lathe in the production. If the test leads to the rejection of the hypothesis, we accept the "*alternative*" $\sigma^2 > \sigma_0^2$ and must use another lathe. In the latter case we say that the test indicates a **significant deviation** of σ^2 from σ_0^2 , that is, a deviation which is not merely caused by the unavoidable influence of chance factors but by the lack of precision of the lathe.

In other cases we may want to compare two things, for example, two different drugs, two methods of performing a certain work, the accuracy of two methods of measurement, the quality of products produced by two different tools, etc. Depending on the result of a suitable test we decide to use one of the two drugs, to introduce the better method of working, etc.

Typical sources for hypotheses are as follows.

1. The hypothesis may come from a quality requirement. The latter may be obtained by producing a large number of items with special care.
2. The values under consideration are known from previous experience.

3. The hypothesis results from a theory to be verified.
4. The hypothesis is a pure guess resulting from occasional observations.

Let us start with a simple introductory example.

Example 1. Observations show that in the random experiment "birth of a single child" the events "birth of a boy" and "birth of a girl" have about the same frequency, and we may test whether the two events have exactly the same probability. Denoting by p the probability of "birth of a boy," our hypothesis to be tested is $p = 50\% = 0.5$. For the test we use a sample of $n = 3000$ children born in Graz in 1962. The number of boys in that sample was 1578.

If the hypothesis is true, we may expect about 1500 boys in a sample of 3000 children. If the number of boys in the sample deviates too much from 1500, this indicates that the hypothesis may be false and we reject it. For this purpose we determine a critical value c (greater than 1500 in our case because we observed more than 1500 boys). If the observed value is greater than c , we reject the hypothesis. If that value is not greater than c , we accept the hypothesis. The basic question now is, how should we choose c , that is, where should we draw the line between small random deviations and large significant deviations. Different people may have different opinions, and to answer the question we must use mathematical arguments. In the present case these are very simple, as follows.

We determine c such that if the hypothesis is true, the probability of observing more than c boys in a sample of 3000 single births is a very small number, call it α . It is customary to choose $\alpha = 1\%$ or 5% . Choosing $\alpha = 1\%$ (or 5%) we risk about once in 100 cases (in 20 cases, respectively) that we reject a hypothesis although it is true. We shall return to this point later. Let us choose $\alpha = 1\%$ and consider the random variable $X = \text{number of boys in 3000 births}$. The critical value c is obtained from the equation

$$(1) \quad P(X > c)_{p=0.5} = \alpha = 0.01,$$

assuming that the hypothesis is true. (This assumption is indicated by the subscript $p = 0.5$.) If the observed value 1578 is greater than c , we reject the hypothesis. If $1578 \leq c$, we accept the hypothesis.

To determine c from (1) we must know the distribution of X . For our purpose the binomial distribution is a sufficiently accurate model. Hence if the hypothesis is true, X has a binomial distribution with $p = 0.5$ and $n = 3000$. This distribution can be approximated by the normal distribution with mean $\mu = np = 1500$ and variance $\sigma^2 = npq = 750$; cf. Sec. 18.10. From (1) we thus obtain

$$P(X > c) = 1 - P(X \leq c) \approx 1 - \Phi\left(\frac{c - 1500}{\sqrt{750}}\right) = 0.01.$$

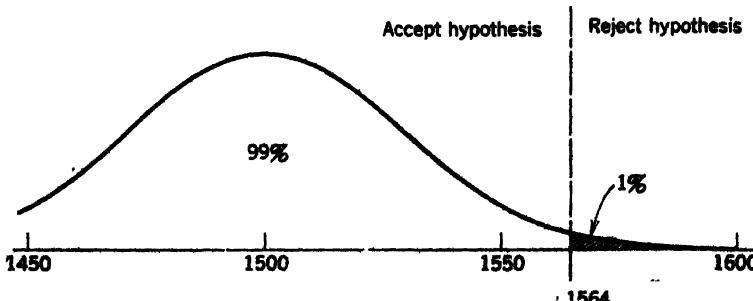


Fig. 409. (Approximate) distribution of X in Ex. 1 if the hypothesis is true.
Critical value $c = 1564$.

Table 50 in Sec. 18.21 yields $(c - 1500)/\sqrt{750} = 2.326$. Hence $c = 1564$. Since $1578 > c$, we reject the hypothesis and assume that $p > 0.5$. This finishes the test.

For a sample of 300 values we would obtain from (1) with $X = \text{number of boys in 300 births}$ the critical value $c = 170$, and a sample in which 158 (the same percentage as in the large sample) are boys would lead to the acceptance of the hypothesis because $158 < c$. This is interesting, because it illustrates that the usefulness of the test increases with increasing sample size n . We should choose n so large that the test will yield information about the question which is of interest in a concrete case. On the other hand, n should not be larger than necessary, for reasons of time and money. In most cases an economical choice of n can be found from a small preliminary experiment.

The hypothesis to be tested is sometimes called the *null hypothesis*, and a counter-assumption (like $p > 0.5$ in Ex. 1) is called an *alternative hypothesis* or, briefly, an *alternative*. The number α (or $100\alpha\%$) is called the *significance level* of the test. c is called the *critical value*. The region containing the values for which we reject the hypothesis is called the *rejection region* or *critical region*. The region of values for which we accept the hypothesis is called the *acceptance region*.

Let u be an unknown parameter in a distribution, and suppose that we want to test the hypothesis $u = u_0$. There are three types of alternatives, namely

- (2) $u > u_0$
- (3) $u < u_0$
- (4) $u \neq u_0$.

(2) and (3) are called *one-sided alternatives*, and (4) is called a *two-sided alternative*. (2) is of the type considered in Ex. 1 (where $u_0 = p = 0.5$ and $u = p > 0.5$); c lies to the right of u_0 , and the rejection region extends from c to ∞ (Fig. 410, upper part). The test is called a *right-sided test*. In the case (3)

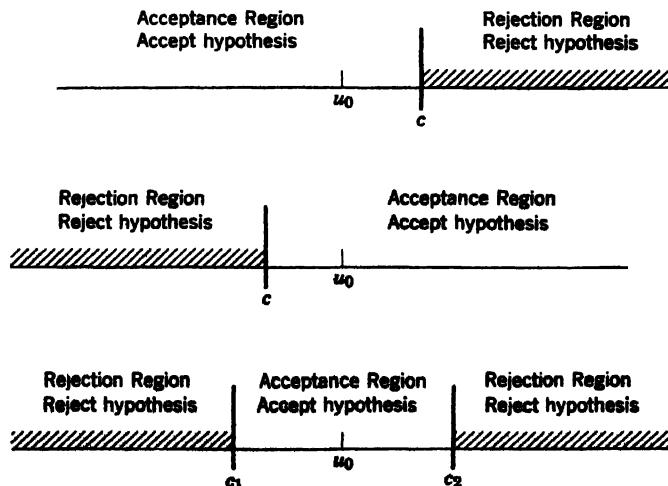


Fig. 410. Test in the case of alternative (2) (upper part of the figure), alternative (3) (middle part), and alternative (4).

the number c lies to the left of u_0 , the rejection region extends from c to $-\infty$ (Fig. 410, middle part), and the test is called a *left-sided test*. Both types of tests are called **one-sided tests**. In the case (4) we have two critical values c_1 and c_2 ($> c_1$), the rejection region extends from c_1 to $-\infty$ and from c_2 to ∞ , and the test is called a **two-sided test**.

All three types of hypotheses are of practical importance. For example, (3) may appear in connection with testing strength of material. u_0 may then be the required strength, and the alternative characterizes an undesirable weakness. The case that the material may be stronger than requested is acceptable, of course, and does therefore not need special attention. (4) may be important, for example, in connection with the diameter of an axle shaft. u_0 then is the required diameter, and slimmer axle-shafts are as bad as thicker ones, so that one has to watch for undesirable deviations from u_0 in both directions.

We shall now consider the risks of making false decisions in a test of a hypothesis $u = u_0$ against an alternative which, for the sake of simplicity, is a single number u_1 . Let $u_1 > u_0$ so that we have a right-sided test, the consideration for a left-sided or a two-sided test being similar. From a given sample x_1, \dots, x_n we compute an estimate $\tilde{u} = g(x_1, \dots, x_n)$, which we may regard as an observed value of the random variable $\tilde{U} = g(X_1, \dots, X_n)$ where x_i is an observed value of X_i . Possibilities of making an error are as follows.

Type I error (cf. Table 41). The hypothesis is true but is rejected because \tilde{U} assumes a value $\tilde{u} > c$. Obviously, the probability of making such an error equals

$$(5) \quad P(\tilde{U} > c)_{u=u_0} = \alpha,$$

the significance level of the test.

Type II error (cf. Table 41). The hypothesis is false, but is accepted because \tilde{U} assumes a value $\tilde{u} \leq c$. The probability of making such an error is denoted by β ; thus

$$(6) \quad P(\tilde{U} \leq c)_{u=u_1} = \beta.$$

$\eta = 1 - \beta$ is called the **power** of the test. Obviously, this is the probability of avoiding a Type II error.

(5) and (6) show that both α and β depend on c , and we would like to choose c such that these probabilities of making errors are as small as possible. But Fig. 411 indicates that these are conflicting requirements, because to let α decrease we must shift c to the right, but then β increases. In practice one first chooses α (5%, 1%, or the like), then determines c , and finally computes β . If β is large so that $\eta = 1 - \beta$ is small, we should repeat the test, choosing a larger sample, for reasons which will appear shortly.

If the alternative is not a single number but is of the form (2)–(4), then η becomes a function of u , which is called the **power function** of the test.

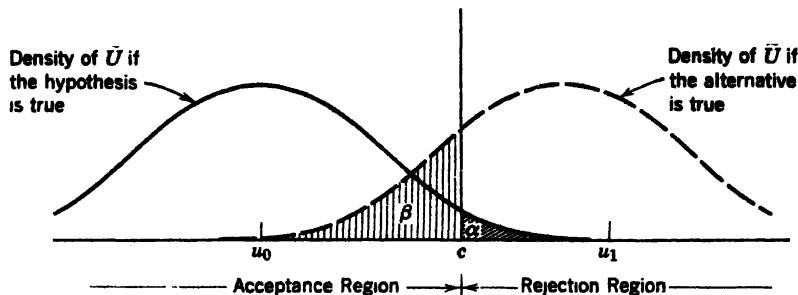


Fig. 411. Illustration of Type I and II errors in testing a hypothesis $u = u_0$ against an alternative $u = u_1 (> u_0)$.

Of course, from a test that leads to the acceptance of a certain hypothesis u_0 , it does *not* follow that this is the only possible hypothesis or the best possible hypothesis. Furthermore, as long as we do not know how large β is, we should be very careful. Perhaps in such a case one should replace the term "accept" by the term "not reject."

Table 411. Type I and Type II Errors in Testing a Hypothesis $u = u_0$ Against an Alternative $u = u_1$.

		Unknown truth	
		$u = u_0$	$u = u_1$
Accepted	$u = u_0$	True decision $P = 1 - \alpha$	Type II error $P = \beta$
	$u = u_1$	Type I error $P = \alpha$	True decision $P = 1 - \beta$

The following examples will explain tests of practically important hypotheses.

Example 2 (Mean of the normal distribution with known variance). Let X be a normal random variable with variance $\sigma^2 = 9$. Using a sample of size $n = 10$ with mean \bar{x} , test the hypothesis $\mu = \mu_0 = 24$ against the three types of alternatives, namely,

$$(a) \mu > \mu_0 \quad (b) \mu < \mu_0 \quad (c) \mu \neq \mu_0.$$

We choose the significance level $\alpha = 0.05$. An estimate of the mean will be obtained from

$$\bar{X} = \frac{1}{n}(X_1 + \cdots + X_n).$$

If the hypothesis is true, \bar{X} is normal with mean $\mu = 24$ and variance $\sigma^2/n = 0.9$, cf. Theorem 2, Sec. 18.14. Hence we may obtain the critical value c from Table 50 in Sec. 18.21.

Case (a). We determine c from $P(\bar{X} > c)_{\mu=24} = \alpha = 0.05$, that is,

$$P(\bar{X} \leq c)_{\mu=24} = \Phi\left(\frac{c - 24}{\sqrt{0.9}}\right) = 1 - \alpha = 0.95.$$

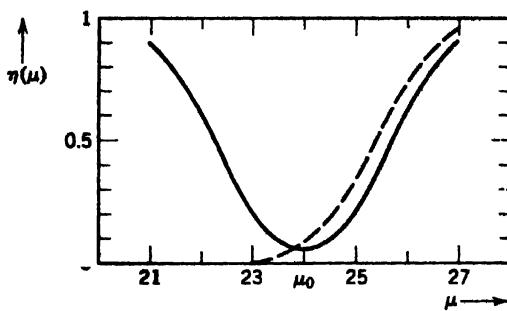


Fig. 412. Power $\eta(\mu)$ in Ex. 2, case (a) (dashed) and case (c).

Table 50, Sec. 18.21, yields $(c - 24)/\sqrt{0.9} = 1.645$, and $c = 25.56$, which is greater than μ_0 , as in the upper part of Fig. 410. If $\bar{x} \leq 25.56$, the hypothesis is accepted. If $\bar{x} > 25.56$, it is rejected. The power of the test is (cf. Fig. 412)

$$(7) \quad \begin{aligned} \eta(\mu) &= P(\bar{X} > 25.56)_{\mu} = 1 - P(\bar{X} \leq 25.56)_{\mu} \\ &= 1 - \Phi\left(\frac{25.56 - \mu}{\sqrt{0.9}}\right) = 1 - \Phi(26.94 - 1.05\mu) \end{aligned}$$

Case (b). The critical value c is obtained from the equation

$$P(\bar{X} \leq c)_{\mu=24} = \Phi\left(\frac{c - 24}{\sqrt{0.9}}\right) = \alpha = 0.05.$$

Table 50, Sec. 18.21, yields $c = 24 - 1.56 = 22.44$. If $\bar{x} \geq 22.44$, we accept the hypothesis. If $\bar{x} < 22.44$, we reject it. The power of the test is

$$(8) \quad \eta(\mu) = P(\bar{X} \leq 22.44)_{\mu} = \Phi\left(\frac{22.44 - \mu}{\sqrt{0.9}}\right) = \Phi(23.65 - 1.05\mu).$$

Case (c). Since the normal distribution is symmetric, we choose c_1 and c_2 equidistant from $\mu = 24$, say, $c_1 = 24 - k$ and $c_2 = 24 + k$, and determine k from

$$P(24 - k \leq \bar{X} \leq 24 + k) = \Phi\left(\frac{k}{\sqrt{0.9}}\right) - \Phi\left(-\frac{k}{\sqrt{0.9}}\right) = 1 - \alpha = 0.95.$$

Using Table 50, Sec. 18.21, we obtain $k/\sqrt{0.9} = 1.960$, $k = 1.86$. Hence $c_1 = 24 - 1.86 = 22.14$ and $c_2 = 24 + 1.86 = 25.86$. If \bar{x} is not smaller than c_1 and not greater than c_2 , we accept the hypothesis. Otherwise we reject it. The power of the test is (cf. Fig. 412)

$$(9) \quad \begin{aligned} \eta(\mu) &= P(\bar{X} < 22.14)_{\mu} + P(\bar{X} > 25.86)_{\mu} = P(\bar{X} < 22.14)_{\mu} + 1 - P(\bar{X} \leq 25.86)_{\mu} \\ &= 1 + \Phi\left(\frac{22.14 - \mu}{\sqrt{0.9}}\right) - \Phi\left(\frac{25.86 - \mu}{\sqrt{0.9}}\right) \\ &= 1 + \Phi(23.34 - 1.05\mu) - \Phi(27.26 - 1.05\mu). \end{aligned}$$

The function $\beta(\mu)$ is called the **operating characteristic (OC)** of the test and its curve the **OC curve**. We know that $\beta(\mu)$ is the probability of making a Type II error (cf. Table 41). In case (c) of Ex. 2 we have (cf. Fig. 413)

$$\beta = \Phi(27.26 - 1.05\mu) - \Phi(23.34 - 1.05\mu).$$

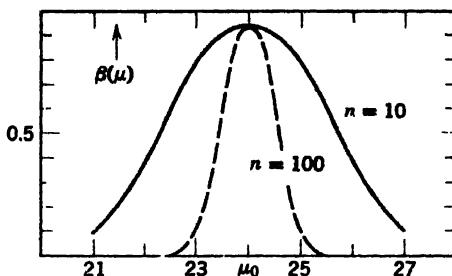


Fig. 413. OC curves in Ex. 2, case (c), for two different sample sizes n .

If we take a larger sample, say, of size $n = 100$ (instead of 10), then $\sigma^2/n = 0.09$ (instead of 0.9), and the critical values are $c_1 = 23.41$ and $c_2 = 24.59$, as can be verified by calculation. Then the operating characteristic of the test in Ex. 2, case (c), is

$$\begin{aligned}\beta(\mu) &= \Phi\left(\frac{24.59 - \mu}{\sqrt{0.09}}\right) - \Phi\left(\frac{23.41 - \mu}{\sqrt{0.09}}\right) \\ &= \Phi(81.97 - 3.33\mu) - \Phi(78.03 - 3.33\mu).\end{aligned}$$

Figure 413 shows that the corresponding OC curve is steeper than that for $n = 10$. This means that the increase of n has led to an improvement of the test. In any practical case n is chosen as small as possible, but so large that the test brings out deviations between μ and μ_0 that are of practical interest. For instance, if in Ex. 2, case (c), deviations of ± 2 units are of interest, then $n = 10$ is much too small (cf. Fig. 413), because when $\mu = 24 - 2 = 22$ or $\mu = 24 + 2 = 26$, then β is almost 50%. On the other hand, $n = 100$ will be sufficient (cf. Fig. 413).

Example 3 (Mean of the normal distribution with unknown variance). The tensile strength of a sample of $n = 16$ manila ropes (diameter 3 in.) was measured. The sample mean was $\bar{x} = 4482$ kg, and the sample standard deviation was $s = 115$ kg (N. C. Wiley, 41st Annual Meeting of the Amer. Soc. for Test. Materials). Assuming that the tensile strength is a normal random variable, test the hypothesis $\mu_0 = 4500$ kg against the alternative $\mu_1 = 4400$ kg. Here μ_0 may be a value given by the manufacturer, while μ_1 may result from previous experience.

We choose the significance level $\alpha = 5\%$. If the hypothesis is true, it follows from Theorem 3, Sec. 18.14, that the random variable

$$T = \sqrt{n} \frac{\bar{X} - \mu_0}{S} = 4 \frac{\bar{X} - 4500}{115}$$

has a t -distribution with $n - 1 = 15$ degrees of freedom. The critical value c is obtained from the equation

$$P(T \leq c)_{\mu_0} = \alpha = 0.05.$$

Table S2 in Sec. 18.21 yields $c = -1.75$. As an observed value of T we obtain from the sample $t = 4(4482 - 4500)/115 = -0.626$. We see that $t > c$ and accept the hypothesis.

Example 4 (Variance of the normal distribution). Using a sample of size $n = 15$ and sample variance $s^2 = 13$ from a normal population, test the hypothesis $\sigma^2 = \sigma_0^2 = 10$ against the alternative $\sigma^2 = \sigma_1^2 = 20$.

We choose the significance level $\alpha = 5\%$. If the hypothesis is true, then

$$Y = (n - 1) \frac{S^2}{\sigma^2} = 14 \frac{S^2}{10} = 1.4 S^2$$

has a chi-square distribution with $n - 1 = 14$ degrees of freedom, cf. Theorem 4, Sec. 18.14. From

$$P(Y > c) = \alpha = 0.05, \quad \text{that is,} \quad P(Y \leq c) = 0.95,$$

and Table 53, Sec. 18.21, with 14 degrees of freedom we obtain $c = 23.68$. This is the critical value of Y . Hence to $S^2 = Y/1.4$ there corresponds the critical value $c^* = 23.68/1.4 = 16.91$. Since $s^2 < c^*$, we accept the hypothesis.

If the alternative is true, the variable

$$Y_1 = 14 \frac{S^2}{\sigma_1^2} = 0.7 S^2$$

has a chi-square distribution with 14 degrees of freedom. Hence our test has the power

$$\eta = P(S^2 > c^*)_{\sigma^2=20} = P(Y_1 > 0.7c^*)_{\sigma^2=20} = 1 - P(Y_1 \leq 11.84)_{\sigma^2=20} \approx 62\%$$

and we see that the Type II risk is very large, namely, 38%. To make that risk smaller, we must increase the sample size.

Example 5 (Comparison of the means of two normal distributions). Using a sample x_1, \dots, x_{n_1} from a normal distribution with unknown mean μ_1 and a sample y_1, \dots, y_{n_2} from another normal distribution with unknown mean μ_2 , we want to test the hypothesis that the means are equal, that is, $\mu_1 = \mu_2$, against an alternative, say, $\mu_1 > \mu_2$. The variances need not be known, but are assumed to be equal.⁵ Two cases are of practical importance:

Case A. The samples have the same size. Furthermore, each value of the first sample corresponds to precisely one value of the other, because corresponding values result from the same person or thing (paired comparisons; for example, two measurements of the same thing by two different methods or two measurements from the two eyes of the same animal, and so on). Then one should form the differences of corresponding values and test the hypothesis that the population corresponding to the differences has the mean 0, using the method in Ex. 3. If we have the choice, this method is better than the following one.

Case B. The two samples are independent and not necessarily of the same size. Then we may proceed, as follows. Suppose that the alternative is $\mu_1 > \mu_2$. We choose a significance level α . We compute the means \bar{x} and \bar{y} and the variances s_1^2 and s_2^2 of the first and second sample, respectively. Using Table 52, Sec. 18.21, with $n_1 + n_2 - 2$ degrees of freedom, we determine c from⁶

$$(10) \quad P(T \leq c) = 1 - \alpha.$$

We finally compute

$$(11) \quad t_0 = \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \frac{\bar{x} - \bar{y}}{\sqrt{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}}.$$

⁵ If the test in the next example shows that the variances differ significantly, then choose two not too small samples of the same size $n_1 = n_2 = n$ (> 30 , say), use the fact that (12) is an observed value of an approximately normal random variable with mean 0 and variance 1, and proceed as in Ex. 2.

⁶ If the alternative is $\mu_1 \neq \mu_2$, then (10) must be replaced by $P(T \leq c_1) = 0.5\alpha$, $P(T \leq c_2) = 1 - 0.5\alpha$.

It can be shown that this is an observed value of a random variable which has a t -distribution with $n_1 + n_2 - 2$ degrees of freedom, provided the hypothesis is true. If $t_0 \leq c$, the hypothesis is accepted. If $t_0 > c$, it is rejected.

Note that for samples of equal size $n_1 = n_2 = n$, formula (11) reduces to

$$(12) \quad t_0 = \sqrt{n} \frac{\bar{x} - \bar{y}}{\sqrt{s_1^2 + s_2^2}}.$$

To illustrate the numerical work, let us consider the two samples

	105	108	86	103	103	107	124	105
and	89	92	84	97	103	107	111	97

showing the relative output of tin plate workers under two different working conditions (J. J. B. WORTH, J. Indust. Eng. 9, 1958, 249–253). We assume the two corresponding populations to be normal and to have the same variance. (The latter assumption will be tested in the next example.) We find $\bar{x} = 105.125$, $\bar{y} = 97.500$, $s_1^2 = 106.125$, $s_2^2 = 84.000$, and $\bar{x} > \bar{y}$ suggests the alternative $\mu_1 > \mu_2$. Choosing the significance level $\alpha = 5\%$, we obtain from (10) and Table 52, Sec. 18.21, with 14 degrees of freedom the value $c = 1.76$. Formula (12) yields $t_0 = \sqrt{8} \cdot 7.625 / \sqrt{190.125} = 1.56 < 1.76$, and we accept the hypothesis $\mu_1 = \mu_2$ that under both conditions the mean output is the same.

Case A applies to the example, because the two first sample values correspond to a certain type of work, the next two were obtained in another kind of work, etc. So we may use the differences

$$16 \quad 16 \quad 2 \quad 6 \quad 0 \quad 0 \quad 13 \quad 8$$

of corresponding sample values and the method in Ex. 3 to test the hypothesis $\mu = 0$ where μ is the mean of the population corresponding to the differences. The sample mean $d = 7.625$ suggests the alternative $\mu > 0$. We choose $\alpha = 5\%$, as before. c is determined from $P(T \leq c)_{\mu=0} = 1 - \alpha = 95\%$ and Table 52, Sec. 18.21, with $n - 1 = 7$ degrees of freedom. We find $c = 1.90$. Then we compute the sample variance $s^2 = 45.696$ and $t = \sqrt{8}(7.625 - 0) / \sqrt{45.696} = 3.19$. Since $3.19 > 1.90$, we reject the hypothesis. Hence our present test, in which we used more information (but the same samples) shows that the difference in output is significant.

Example 6 (Comparison of the variances of two normal distributions). Using the two samples in the last example, test the hypothesis $\sigma_1^2 = \sigma_2^2$. We find $s_1^2 = 106.125$, $s_2^2 = 84.000$, and this suggests the alternative $\sigma_1^2 > \sigma_2^2$. We choose the significance level $\alpha = 5\%$. Using the equation $P(V \leq c) = 1 - \alpha = 95\%$ and Table 54, Sec. 18.21, with $(n_1 - 1, n_2 - 1) = (7, 7)$ degrees of freedom, we determine $c = 3.79$. We finally compute $v_0 = s_1^2 / s_2^2 = 1.26$. Since $v_0 \leq c$, we accept the hypothesis. If $v_0 > c$, we would reject it.

This test is justified by the fact that t_0 is an observed value of a random variable which has a so-called **F-distribution** with $(n_1 - 1, n_2 - 1)$ degrees of freedom, provided the hypothesis is true. (Proof in Ref. [13], cf. Appendix 1.) The F-distribution with (m, n) degrees of freedom was introduced by R. A. Fisher and has the distribution function $F(z) = 0$ when $z < 0$ and

$$(13) \quad F(z) = K_{mn} \int_0^z t^{(m-2)/2} (mt + n)^{-(m+n)/2} dt \quad (z \geq 0)$$

where $K_{mn} = m^{m/2} n^{n/2} \Gamma(\frac{1}{2}m + \frac{1}{2}n) / \Gamma(\frac{1}{2}m) \Gamma(\frac{1}{2}n)$.

Problems

- Using the data of Buffon in Table 35, Sec. 18.5, test the hypothesis that the coin is fair, that is, that heads and tails have the same probability of occurrence, against the alternative that heads are more likely than tails.

2. Same test as in Prob. 1, using the data of Pearson in Table 35, Sec. 18.5.
3. Suppose that the sample in Ex. 1, Sec. 18.3, was taken from a normal population. Test the hypothesis $\mu = 0.80$ against the alternative $\mu > 0.80$.
4. Assuming that the sample in Prob. 17, Sec. 18.14, was taken from a normal population, test the hypothesis $\mu = 1.2$ against the alternative $\mu < 1.2$.
5. In Prob. 4, test the hypothesis (a) $\mu = 1.25$, (b) $\mu = 1.15$ against a sensible alternative.
6. The starch content of 16 potatoes was measured by two different methods. The differences [%] of corresponding values were

2 0 0 1 2 2 3 -3 1 2 3 0 -1 1 -2 1.

Assuming that the corresponding population is normal, test the hypothesis that there is no significant difference between the two methods.

7. Graph the OC curves in Ex. 2, cases (a) and (b).
8. For checking the reliability of a certain electrical method for measuring temperatures the following two independent samples were obtained [C].

106.9	106.3	107.0	106.0	104.9
106.5	106.7	106.8	106.1	105.6

Assuming that the corresponding populations are normal and have the same variance, test the hypothesis that the means do not differ significantly.

9. For comparing two methods of hand-picking (used in industry for eliminating bad peanuts, beans, etc.) 9 workers picked black beans from a given lot of white and black beans. The numbers of beans picked during an interval of time of a certain length were as follows.

Worker No.	1	2	3	4	5	6	7	8	9
Method A	214	253	276	215	238	221	210	229	269
Method B	281	279	260	230	267	253	205	265	299

Assuming that the corresponding populations are normal and have the same variance, test the hypothesis that there is no significant difference between the effectiveness of the two methods.

10. The following two samples (weight of dust in certain tubes, measured in mg) were obtained in an experiment by the British Coal Utilisation Research Association:

75	20	70	70	85	90	100	40	35	65	90	35
20	35	55	50	65	40						

Assuming that the corresponding populations are normal, find whether the variances of these populations differ significantly from each other.

18.16 Quality Control

No process of production is so perfect that all the products are completely alike. There is always a small variation that is caused by a great number of small uncontrollable influences and must therefore be regarded as a chance variation. It is important to make sure that the products have required values

(for example, length, strength, or what other property may be of importance in a particular case). For this purpose one makes a test of the hypothesis that the products have the required property, say, $\mu = \mu_0$, where μ_0 is a required value. If this is done after the production of a whole lot (for example 100,000 screws) is completed, the test will tell us how good or how bad the products are, but there is no way of changing undesirable results or facts. It is much better to test while the production is still in progress. This is done at regular intervals of time (every 30 minutes, every hour or so) and is called **quality control**. At each time the sample size is the same, in practice 3 to 10 items. If the hypothesis is rejected, we stop the process of production and look for the trouble that causes the deviation.

If we stop the production process although it is progressing properly, we make a Type I error. If we do not stop the process although something is not in order, we make a Type II error (cf. Sec. 18.15).

The result of each test is marked in graphical form on what is called a **control chart**. This makes quality control particularly effective. An example for the control of the mean is shown in Fig. 414. The **control limits** c_1 and c_2 in this figure correspond to the critical values c_1 and c_2 in Fig. 410 and in Ex. 2, Sec. 18.15, case (c). The number c_1 is called the *lower control limit* (LCL), and c_2 is called the *upper control limit* (UCL). As soon as a sample mean lies outside the range between the control limits, the hypothesis is rejected and the process of production is stopped.

If we choose control limits that are too narrow, we risk interrupting the production process unnecessarily often although it runs properly. The usual significance level is $\sigma = 1\%$. From Theorem 2, Sec. 18.14, and Table 49,

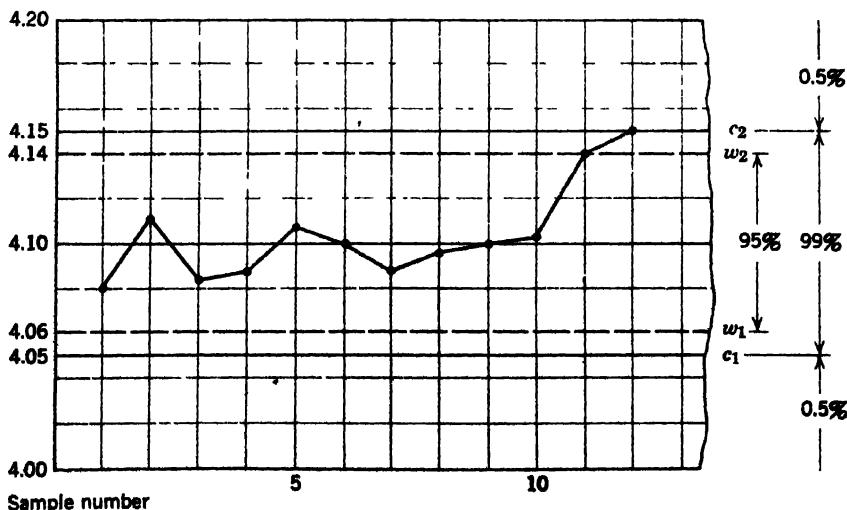


Fig. 414. Control chart for the mean value (diameter of a hole in a plate) where $\mu_0 = 4.10$ mm, $c_1 = 4.05$ mm, $c_2 = 4.15$ mm, $w_1 = 4.06$ mm, $w_2 = 4.14$ mm.

Sec. 18.21, we see that in the case of the normal distribution the corresponding control limits are

$$(1) \quad c_1 = \mu_0 - 2.58 \frac{\sigma}{\sqrt{n}} \quad \text{and} \quad c_2 = \mu_0 + 2.58 \frac{\sigma}{\sqrt{n}}.$$

Here σ is assumed to be known. If σ is unknown, we may calculate the standard deviations of the first 20 or 30 samples and take their arithmetic mean as an approximation of σ .

Sometimes control charts also show the limits

$$(2) \quad w_1 = \mu_0 - 1.96 \frac{\sigma}{\sqrt{n}} \quad \text{and} \quad w_2 = \mu_0 + 1.96 \frac{\sigma}{\sqrt{n}}$$

corresponding to the significance level $\alpha = 5\%$. If a sample mean lies between w_1 and c_1 or between w_2 and c_2 , then another sample is taken right away and the test is repeated.

What usually happens in practice when quality control charts are introduced is this: After establishing control limits on the basis of 20 or more samples, and after searching for and eliminating causes for process trouble every time a point falls outside the control limits, it is soon found that the variability from mean to mean becomes smaller than it was for the initial set of samples. A new set of samples can then be taken and new upper and lower control limits established. These control limits, in practice, will usually be closer together than the original ones. After a few stages of this kind we arrive at a stage of statistical control involving control limits which are about as close together as can be hoped for without revolutionary changes in the manufacturing process.

Table 42. Sample Values in Fig. 414.

Sample Number	Sample Values					\bar{x}	s	R
1	4.06	4.08	4.08	4.08	4.10	4.080	0.014	0.04
2	4.10	4.10	4.12	4.12	4.12	4.112	0.011	0.02
3	4.06	4.06	4.08	4.10	4.12	4.084	0.026	0.06
4	4.06	4.08	4.08	4.10	4.12	4.088	0.023	0.06
5	4.08	4.10	4.12	4.12	4.12	4.108	0.018	0.04
6	4.08	4.10	4.10	4.10	4.12	4.100	0.014	0.04
7	4.06	4.08	4.08	4.10	4.12	4.088	0.023	0.06
8	4.08	4.08	4.10	4.10	4.12	4.096	0.017	0.04
9	4.06	4.08	4.10	4.12	4.14	4.100	0.032	0.08
10	4.06	4.08	4.10	4.12	4.16	4.104	0.038	0.10
11	4.12	4.14	4.14	4.14	4.16	4.140	0.014	0.04
12	4.14	4.14	4.16	4.16	4.16	4.152	0.011	0.02

(C. H. FISCHER, Telefunken-Zeitung 36, 1963, 43)

Sometimes we control the variance, in addition to the mean. The corresponding control chart shows a control limit corresponding to the number c^* in Ex. 4, Sec. 18.15, with $\alpha = 1\%$ instead of 5%. It is customary to print the two control charts for the mean and the variance on the same sheet one below the other.

If we control σ^2 , we must compute s^2 . Since this may be difficult for mathematically untrained persons, it is desirable to replace the control of the variance by that of the range R ($=$ largest sample value minus smallest sample value). It can be shown that in the case of the normal distribution the expectation of the random variable R^* for which R is an observed value, is proportional to σ , say, $E(R^*) = \lambda_n \sigma$. The factor of proportionality λ_n depends on n and has the values:

n	2	3	4	5	6	7	8	9	10
λ_n	1.1	1.7	2.1	2.3	2.5	2.7	2.9	3.0	3.1

s is a better estimate of σ than R and should be used in particular if single measurements are expensive, because R does not use all the information contained in the sample, but depends on two sample values only.

Problems

- Suppose a machine for filling cans with lubricating oil is set so that it will generate fillings which form a normal population with mean 1 gal and standard deviation 0.02 gal. Set up a control chart of the type shown in Fig. 414 for controlling the mean (that is, find c_1 , c_2 , w_1 , and w_2), assuming that the sample size is 4.
- What values c_1 and c_2 do we obtain in Prob. 1 if we require the significance level $\alpha = 0.3\%$? Hint: cf. (7) in Sec. 18.10.
- How does the meaning of the control limits (1) change if we apply a control chart with these limits in the case of a population which is not normal?
- What values c_1 and c_2 should we use instead of (1) if we use the sum $x_1 + \dots + x_n$ of the sample values instead of \bar{x} ?
- Determine the values c_1 and c_2 mentioned in Prob. 4 in the case of Fig. 414.

18.17 Acceptance Sampling

Acceptance sampling is applied in mass production when a *producer* is supplying to a *consumer* lots of N items. In such a situation the decision to accept or reject an individual lot must be made. This decision is often based on the result of inspecting a sample of size n from the lot. If the number of defectives x in the sample is not greater than a specified number c ($< n$), the lot is accepted. If $x > c$, the lot is rejected. c is called the *allowable number of defectives*. It is clear that the producer and the consumer must agree on a certain *sampling plan*, that is, on a certain sample size n and an allowable number c . Such a plan is called a *single sampling plan* because it is based on a

single sample. *Double sampling plans*, in which two samples are used, will be mentioned later.

Let A be the event that a lot is accepted. It is clear that the corresponding probability $P(A)$ depends not only on n and c but also on the number of defectives in the lot. Let M denote this number, let the random variable X be the number of defectives in the sample, and suppose that we sample without replacement. Then (cf. Sec. 18.9)

$$(1) \quad P(A) = P(X \leq c) = \sum_{x=0}^c \binom{M}{x} \binom{N-M}{n-x} / \binom{N}{n}.$$

If $M = 0$ (no defectives in the lot), then X must assume the value 0, and

$$P(A) = \binom{0}{0} \binom{N}{n} / \binom{N}{n} = 1.$$

For fixed n and c and increasing M the probability $P(A)$ decreases. If $M = n$ (all items of the lot defective), then X must assume the value n , and $P(A) = P(X \leq c) = 0$ because $c < n$.

The ratio $\theta = M/N$ is called the *fraction defective* in the lot. Note that $M = N\theta$, and (1) may be written

$$(2) \quad P(A; \theta) = \sum_{x=0}^c \binom{N\theta}{x} \binom{N-N\theta}{n-x} / \binom{N}{n}.$$

Since θ can have one of the $N+1$ values $0, 1/N, 2/N, \dots, N/N$, the probability $P(A)$ is defined for these values only. For fixed n and c we may plot $P(A)$ as a function of θ . These are $N+1$ points. Through these points we may then draw a smooth curve, which is called the *operating characteristic curve (OC curve)* of the sampling plan considered.

Example 1. Suppose that certain tool bits are packaged 20 to a box, and the following single sampling plan is used: A sample of 2 tool bits is drawn, and the corresponding box is accepted if and only if both bits in the sample are good. In this case, $N = 20$, $n = 2$, $c = 0$, and (2) takes the form

$$P(A; \theta) = \binom{20\theta}{0} \binom{20-20\theta}{2} / \binom{20}{2} = \frac{(20-20\theta)(19-20\theta)}{380}.$$

The numerical values are

θ	0.00	0.05	0.10	0.15	0.20	...
$P(A; \theta)$	1.00	0.90	0.81	0.72	0.63	...

The OC curve is shown in Fig. 415.

In most practical cases θ will be small (less than 10%). In many cases the lot size N will be very large (1000, 10,000, etc.), so that we may approximate the hypergeometric distribution in (1) and (2) by the binomial distribution with $p = \theta$. Then if n is such that $n\theta$ is moderate (say, less than 20), we may approximate that distribution by the Poisson distribution with mean μ

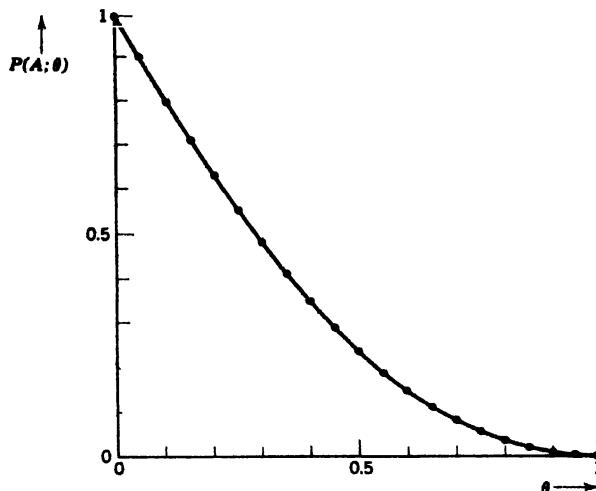


Fig. 415. OC curve of the single sampling plan with $n = 2$ and $c = 0$ for lots of size $N = 20$.

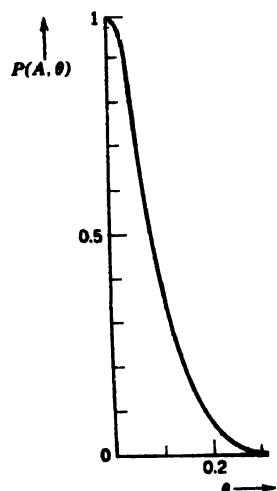


Fig. 416. OC curve in Example 2.

From (2) we then have

$$(3) \quad P(A; \theta) \sim e^{-\mu} \sum_{x=0}^c \frac{\mu^x}{x!} \quad (\mu = n\theta).$$

Example 2. Suppose that for large lots the following single sampling plan is used. A sample of size $n = 20$ is taken. If it contains not more than 1 defective, the lot is accepted. If the sample contains 2 or more defectives, the lot is rejected. In this plan, we obtain from (3)

$$P(A; \theta) \sim e^{-20\theta}(1 + 20\theta).$$

The corresponding OC curve is shown in Fig. 416.

We shall now consider the two types of possible errors in acceptance sampling and the related problem of choosing n and c . In acceptance sampling the producer and the consumer have different interests. The producer may require the probability of rejecting a "good" or "acceptable" lot to be a small number, call it α . The consumer may demand the probability of accepting a "bad" or "unacceptable" lot to be a small number β . More precisely, suppose that the two parties agree that a lot for which θ does not exceed a certain number θ_0 is an *acceptable lot* while a lot for which θ is greater or equal to a certain number θ_1 is an *unacceptable lot*. Then α is the probability of rejecting a lot with $\theta \leq \theta_0$ and is called *producer's risk*. This corresponds to a Type I error in testing a hypothesis (Sec. 18.15). β is the probability of accepting a lot with $\theta \geq \theta_1$ and is called *consumer's risk*. This corresponds to a Type II error in Sec. 18.15. Figure 417 shows an illustrative example. θ_0 is called the *acceptable quality level (AQL)*, and θ_1 is called the *lot tolerance per cent defective (LTPD)*. A lot with $\theta_0 < \theta < \theta_1$ may be called an *indifferent lot*.

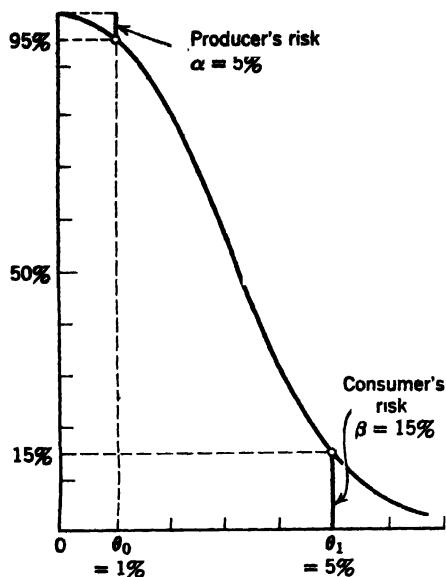


Fig. 417. OC curve, producer's and consumer's risk.

From Fig. 417 we see that the points $(\theta_0, 1 - \alpha)$ and (θ_1, β) lie on the OC curve. It can be shown that for large lots we can choose θ_0 , $\theta_1 (> \theta_0)$, α , β and then determine n and c such that the OC curve passes through those prescribed points. Sampling plans for specified α , β , θ_0 , and θ_1 have been published; cf. Ref. [J3], Appendix 1.

There is a close relationship between sampling inspection and testing a hypothesis, as follows:

Sampling Inspection

Acceptable quality level (<i>AQL</i>) $\theta = \theta_0$
Lot tolerance per cent defectives (<i>LTPD</i>)
$\theta = \theta_1$
Allowable number of defectives c
Producer's risk α of rejecting a lot with $\theta \leq \theta_0$
Consumer's risk β of accepting a lot with $\theta \geq \theta_1$

Hypothesis Testing

Hypothesis $\theta = \theta_0$
Alternative $\theta = \theta_1$
Critical value c
Probability α of making a Type I error (significance level)
Probability β of making a Type II error

The sampling procedure by itself does not protect the consumer sufficiently well. In fact, if the producer is permitted to resubmit a rejected lot without telling that the lot has already been rejected, then even bad lots will eventually be accepted. To protect the consumer against this and other possibilities, the producer may agree with the consumer that a rejected lot is *rectified*, that is, is inspected 100%, item by item, and all defective items in the lot are

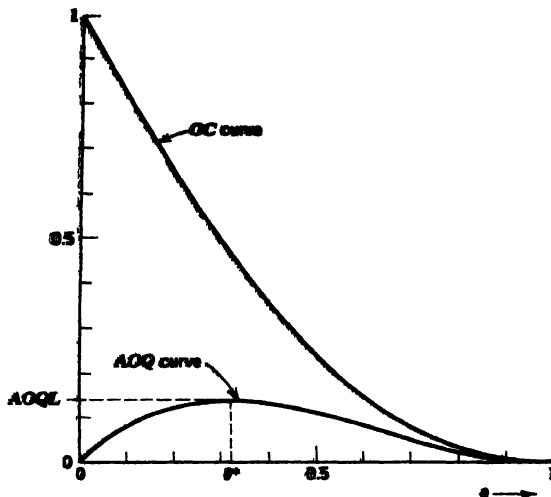


Fig. 418. *OC* curve and *AOQ* curve for the sampling plan in Fig. 415.

removed and replaced by nondefective items.⁷ Suppose that a plant produces 100θ% defectives and rejected lots are rectified. Then K lots of size N contain KN items, $KN\theta$ of which are defective. $KP(A; \theta)$ of the lots are accepted; these contain a total of $KPN\theta$ defective items. The rejected and rectified lots contain no defectives. Hence after the rectification the fraction defective in the K lots equals $KPN\theta/KN = \theta P(A; \theta)$. This function of θ is called the *average outgoing quality* (*AOQ*) and is denoted by $AOQ(\theta)$. Thus

$$(4) \quad AOQ(\theta) = \theta P(A; \theta).$$

A sampling plan being given, this function and its graph, the *average outgoing quality curve* (*AOQ* curve), can readily be obtained from $P(A; \theta)$ and the *OC* curve. An example is shown in Fig. 418.

Clearly $AOQ(0) = 0$. Also, $AOQ(1) = 0$ because $P(A; 1) = 0$. Since $AOQ \geq 0$, we conclude that this function must have a maximum at some $\theta = \theta^*$. The corresponding value $AOQ(\theta^*)$ is called the *average outgoing quality limit* (*AOQL*). This is the worst average quality which may be expected to be accepted under the rectifying procedure.

It turns out that several single sampling plans may correspond to the same *AOQL*; cf. Ref. [J3] in Appendix 1. Hence if the *AOQL* is all the consumer cares about, the producer has some freedom in choosing a sampling plan and may select a plan which minimizes the amount of sampling, that is, the number of inspected items per lot. This number is

$$n P(A; \theta) + N(1 - P(A; \theta))$$

⁷ Of course, rectification is impossible if the inspection is destructive, or is not worthwhile if it is too expensive, compared with the value of the lot. The rejected lot may then be sold at a cut-rate price or scrapped.

where the first term corresponds to the accepted lots and the last term to the rejected and rectified lots; in fact, rectification requires the inspection of all N items of the lot, and $1 - P(A; \theta)$ is the probability of rejecting a lot.

We mention that inspection work may be saved by using a **double sampling plan**, in which the sample of size n is broken up into two samples of sizes n_1 and n_2 (where $n_1 + n_2 = n$). If the lot is very good or very bad, it may then be possible to decide about acceptance or rejectance, using one sample only, so that the other sample will be necessary in the case of a lot of intermediate quality only. Ref. [J3] contains double sampling plans using rectifying inspection of the following type (where x_1 and x_2 are the numbers of defectives in those two samples).

- (1) If $x_1 \leq c_1$, accept the lot. If $x_1 > c_2$, reject the lot.
- (2) If $c_1 < x_1 \leq c_2$, use the second sample, too. If $x_1 + x_2 \leq c_2$, accept the lot. If $x_1 + x_2 > c_2$, reject the lot.

In the sampling plans considered the decision of accepting or rejecting a lot is based on the number of defective items in a sample. Such a plan is therefore called a *sampling plan based on attributes*. In such a plan the only information about an inspected item is whether the item is defective or not. For example, in acceptance sampling of lots of air mail envelopes we may consider the weight X of an envelope and regard an envelope as defective if its weight exceeds a certain limit, say 2.1 g. In this procedure we base the decision on the number of defectives in a sample. However we may arrive at a more sensitive sampling plan based on the mean \bar{x} of a sample of size n . The lot is then accepted if \bar{x} does not exceed some number c^* . It is rejected if $\bar{x} > c^*$.

In this sampling we may choose n and c^* such that the producer's and consumer's risks have prescribed values α and β , respectively. For this purpose we must know the distribution of the above variable X . Suppose that X is normal with mean μ and variance σ^2 . Then \bar{x} is an observed value of a random variable \bar{X} which is normal with mean μ and variance σ^2/n ; cf. Theorem 2, Sec. 18.14. We may now determine n and c^* from

$$P(\bar{X} \leq c^*)_{\mu_0} = \Phi\left(\sqrt{n} \frac{c^* - \mu_0}{\sigma}\right) = 1 - \alpha$$

$$P(\bar{X} \leq c^*)_{\mu_1} = \Phi\left(\sqrt{n} \frac{c^* - \mu_1}{\sigma}\right) = \beta.$$

For example, suppose that $\sigma = 0.05$, and we request $\alpha = 5\%$, $\beta = 10\%$. $\mu_0 = 1.98$, $\mu_1 = 2.00$. Then $1 - \alpha = 0.95$, and Table 50 yields

$$\sqrt{n} \frac{c^* - 1.98}{0.05} = 1.645, \quad \sqrt{n} \frac{c^* - 2.00}{0.05} = -1.282.$$

By forming the quotient of the two equations we find

$$\frac{c^* - 1.98}{c^* - 2.00} = - \frac{1.645}{1.282}; \quad \text{thus} \quad c^* = 1.991 \ 24.$$

By inserting this into one of the equations we obtain $n = 53.5 \approx 54$.

Problems

1. Find the binomial approximation of the hypergeometric distribution in Ex. 1 and compare the approximate and the accurate values.
2. Using the Poisson approximation of the hypergeometric distribution, graph the *OC* curve corresponding to the following single sampling plan for large lots. A sample of $n = 30$ items is drawn from the lot and the 30 items are inspected. The lot is accepted if the sample contains not more than $c = 1$ defective. Otherwise the lot is rejected.
3. Graph the *AOQ* curve in Prob. 2. Determine the *AOQL*, assuming that rectification is possible.
4. Do the work required in Prob. 2 when $n = 50$ and $c = 0$.
5. Graph the *OC* curve and the *AOQ* curve for the single sampling plan for large lots with $n = 5$ and $c = 0$.
6. Same task as in Prob. 5 when $n = 4$ and $c = 1$.
7. Estimate θ_0 in Prob. 5 if the producer's risk is 5%. Estimate θ_1 in Prob. 5 if the consumer's risk is 10%.

18.18 Goodness of Fit. χ^2 -Test

Using a sample x_1, \dots, x_n , we want to test the hypothesis that a certain function $F(x)$ is the distribution function of the population from which the sample was taken. Clearly, the sample distribution function $\tilde{F}(x)$ is an approximation of $F(x)$, and if it approximates $F(x)$ "sufficiently well," we shall accept the hypothesis. On the other hand, if $\tilde{F}(x)$ deviates "too much" from $F(x)$, we shall reject the hypothesis. For this purpose we must know how much $\tilde{F}(x)$ can differ from $F(x)$ if the hypothesis is true, that is, we have to introduce a measure for the deviation of $\tilde{F}(x)$ from $F(x)$ and we must know the probability distribution of this measure under the assumption that the hypothesis is true. Then we can determine a number c such that if the hypothesis is true, a deviation greater than c has a small preassigned probability α , which is called the *significance level* of the test. If nevertheless such a large deviation occurs, we have a reason for doubting whether the hypothesis is true and we reject it. On the other hand, if the deviation does not exceed c so that $\tilde{F}(x)$ approximates $F(x)$ sufficiently well, we accept the hypothesis. Of course, the acceptance of a hypothesis does not exclude that there may be other functions $F(x)$ which would also be accepted in the test; in this respect the situation is similar to that in Sec. 18.15.

Table 43. Chi-square Test for
of a Which α That $F(x)$ is the Distribution
 x_1, \dots, x_n is Taken.

1st step. Subdivide the x -axis into K intervals I_1, I_2, \dots, I_K such that each interval contains at least 5 values of the given sample x_1, \dots, x_n . Determine the number b_j of sample values in the interval I_j ($j = 1, \dots, K$). If a sample value lies at a common boundary point of two intervals, add 0.5 to each of the two corresponding b_j .

2nd step. Using $F(x)$, compute the probability p_j that the random variable X under consideration assumes any value in the interval I_j ($j = 1, \dots, K$). Compute

$$e_j = np_j.$$

(This is the number of sample values theoretically expected in I_j if the hypothesis is true.)

3rd step. Compute the deviation

$$(1) \quad \chi^2 = \sum_{j=1}^K \frac{(b_j - e_j)^2}{e_j}.$$

4th step. Choose a significance level (5%, 1%, or the like).

5th step. Determine the solution c of the equation

$$P(\chi^2 \leq c) = 1 - \alpha$$

from the table of the chi-square distribution with $K - 1$ degrees of freedom (Table 53, Sec. 18.21). If r parameters of $F(x)$ are unknown and their maximum likelihood estimates (Sec. 18.13) are used, then use $K - r - 1$ degrees of freedom (instead of $K - 1$). If $\chi^2 \leq c$, the hypothesis is accepted. If $\chi^2 > c$, it is rejected.

Table 44. Calculations in Example 1.

x_j	$\frac{x_j - 364.7}{26.7}$	$\Phi\left(\frac{x_j - 364.7}{26.7}\right)$	$e_j = 100p_j$	b_j	Term in (1)
$-\infty \dots 325$	$-\infty \dots -1.49$	$0.0000 \dots 0.0681$	6.81	6	0.096
$325 \dots 335$	$-1.49 \dots -1.11$	$0.0681 \dots 0.1335$	6.54	6	0.045
$335 \dots 345$	$-1.11 \dots -0.74$	$0.1335 \dots 0.2296$	9.61	11	0.291
$345 \dots 355$	$-0.74 \dots -0.36$	$0.2296 \dots 0.3594$	12.98	14	0.660
$355 \dots 365$	$-0.36 \dots 0.80$	$0.3594 \dots 0.5000$	14.06	16	0.268
$365 \dots 375$	$0.80 \dots 0.39$	$0.5000 \dots 0.6517$	15.17	15	0.992
$375 \dots 385$	$0.39 \dots 0.76$	$0.6517 \dots 0.7764$	12.47	8	1.682
$385 \dots 395$	$0.76 \dots 1.13$	$0.7764 \dots 0.8708$	9.44	10	0.833
$395 \dots 405$	$1.13 \dots 1.51$	$0.8708 \dots 0.9345$	6.37	8	0.417
$405 \dots \infty$	$1.51 \dots \infty$	$0.9345 \dots 1.0000$	6.55	6	0.046

$$\chi^2 = 2.790$$

Table 43 shows a test of this type, which was introduced by R. A. Fisher. This test is justified by the fact that if the hypothesis is true, then χ_0^2 is an observed value of a random variable whose distribution function approaches that of the chi-square distribution with $K - 1$ degrees of freedom (or $K - r - 1$ degrees of freedom if r parameters are estimated) as n approaches infinity. The requirement that at least 5 sample values should lie in each interval in Table 43 results from the fact that that random variable is only asymptotically normal. A proof can be found in Ref. [I3], cf. Appendix 1.

Example 1. Test whether the population from which the sample in Table 28 was taken is normal. The maximum likelihood estimates for μ and σ are $\bar{x} = 364.7$ and $\tilde{s}^2 = (99/100)s^2 = 712.9$ (cf. Table 34 in Sec. 18.3). The calculation in Table 44 yields $\chi_0^2 = 2.790$. We choose $\alpha = 5\%$. Since $K = 10$ and we estimated $r = 2$ parameters, we have to use Table 53, Sec. 18.21, with $K - r - 1 = 7$ degrees of freedom, finding $c = 14.07$ as the solution of $P(\chi^2 \leq c) = 95\%$. Since $\chi_0^2 < c$, we accept the hypothesis that that population is normal.

Problems

- In one of G. Mendel's classical experiments the result was 355 yellow peas and 123 green peas. Is this in agreement with Mendel's theory according to which in the present case the ratio yellow peas:green peas should be 3:1?
- A coin was tossed 50 times. Head was obtained 27 times. Using this result, test the hypothesis that heads and tails have the same probability.
- In the sample in Prob. 2, what is the minimum number of heads (greater than 25) that would lead to the rejection of the hypothesis if $\alpha = 5\%$ is used?
- The following sample consists of values of the mass of cotton material (measured in gram/m²). Test the hypothesis that the corresponding population is normal.

Class midpoint	96	98	100	102	104	106	108	110	112
Absolute class frequency	1	0	1	2	8	19	28	30	41
Class midpoint	114	116	118	120	122	124	126	128	130
Absolute class frequency	66	50	27	8	5	3	0	1	1

- Is the sample in Table 30, Sec. 18.2, taken from a normal population?

18.19 Nonparametric Tests

The tests in Sec. 18.15 refer to normal populations. In many applications the populations will not be normal. Then we may apply a so-called *nonparametric test*, which is based on order statistics and therefore is valid for any *continuous* distribution. These tests are simple. However, for a normal

distribution the tests in Sec. 18.15 yield better results. To illustrate the basic idea of a nonparametric test, let us consider two typical examples.

Example 1 (Sign test). Using the sample of differences in Ex. 5, Sec. 18.15, that is,

$$16 \quad 16 \quad 2 \quad 6 \quad 0 \quad 0 \quad 13 \quad 8,$$

test the hypothesis $\mu = 0$, that is, the outputs corresponding to the two different working conditions do not differ significantly.

We choose the significance level $\alpha = 5\%$. If the hypothesis is true, the probability p of a positive difference is the same as that of a negative difference. Hence $p = 0.5$, and the random variable

$$X = \text{number of positive values among } n \text{ values}$$

has a binomial distribution with $p = 0.5$. Our sample has 8 values. We omit the values 0 which do not contribute to the decision. Then 6 values are left, all of which are positive. Since

$$P(X = 6) = \binom{6}{6} (0.5)^6 (0.5)^0 = 0.0156 = 1.56\% < \alpha,$$

we reject the hypothesis and accept the alternative $\mu > 0$ suggested by $d > 0$ (cf. Ex. 5 in Sec. 18.15)

If 1 of those 6 values would be negative, we would have

$$P(X \geq 5) = \binom{6}{5} (0.5)^5 \cdot 0.5 + \binom{6}{6} (0.5)^6 = 10.9\%$$

and we would accept the hypothesis. If in a sample the number of negative values is greater than that of the positive values, we take the random variable

$$X = \text{number of negative values among } n \text{ values}$$

and proceed as before. (Values of the binomial distribution are contained in Table 47, Sec. 18.21.)

Example 2 (Test for arbitrary trend). A certain machine is used for cutting lengths of wire. Five successive pieces had the lengths

$$29 \quad 31 \quad 28 \quad 30 \quad 32.$$

Using this sample, test the hypothesis that there is *no trend*, that is, the machine does not have the tendency of producing longer and longer pieces. The sample suggests the alternative that there is *positive trend*, that is, there is the tendency that successive pieces get longer.

We count the number of *inversions* in the sample, that is, the number of times a larger value precedes a smaller value:

$$\begin{array}{ll} 29 \text{ precedes } 28 & (1 \text{ inversion}), \\ 31 \text{ precedes } 28 \text{ and } 30 & (2 \text{ inversions}). \end{array}$$

The remaining three sample values follow in ascending order. Hence in the sample there are $1 + 2 = 3$ inversions.

We now consider the random variable

$$I = \text{number of inversions}.$$

If the hypothesis is true (no trend), then each of the $5! = 120$ permutations of 5 elements 1 2 3 4 5 has the same probability (1/120). We arrange these permutations according to their number of inversions:

$I = 0$	$I = 1$	$I = 2$	$I = 3$	
1 2 3 4 5	1 2 3 5 4	1 2 4 5 3	1 2 5 4 3	1 2 5 4 3
	1 2 4 3 5	1 2 5 3 4	1 3 4 5 2	
	1 3 2 4 5	1 3 2 5 4	1 3 5 2 4	
	2 1 3 4 5	1 3 4 2 5	1 4 2 5 3	
		1 4 2 3 5	1 4 3 2 5	
		2 1 3 5 4	1 5 2 3 4	
		2 1 4 3 5	2 1 4 5 3	
		2 3 1 4 5	2 1 5 3 4	etc.
		3 1 2 4 5	2 3 1 5 4	
			2 3 4 1 5	
			2 4 1 3 5	
			3 1 2 5 4	
			3 1 4 2 5	
			3 2 1 4 5	
			4 1 2 3 5	

From this we obtain

$$P(I \leq 3) = \frac{1}{120} + \frac{4}{120} + \frac{8}{120} + \frac{15}{120} = \frac{38}{120} = 24\%.$$

Hence we accept the hypothesis.

Values of the distribution function of I in the case of no trend are shown in Table 55, Sec. 18.21. Our method and those values refer to *continuous* distributions. Theoretically we may then expect that all the values of a sample are different. Practically, some sample values may still be equal, because of rounding off. If m values are equal, add $m(m - 1)/4$ (= mean value of the inversions in the case of the permutations of m elements), that is, $\frac{1}{2}$ for each pair of equal values, $\frac{3}{2}$ for each triple, etc. If the population has a discrete distribution, we must modify the method, taking into account that theoretically each number may occur more than once in a sample.

Problems

1. Apply the sign test to Prob. 6, Sec. 18.15. (The necessary values of the binomial distribution are not contained in Table 47, Sec. 18.21, but must be computed.)
2. Each of 10 patients were given two different types of sleeping pills *A* and *B*. The following table shows the effect (extension of sleep measured in hours).

<i>A</i>	1.9	0.8	1.1	0.1	-0.1	4.4	5.5	1.6	4.6	3.4
<i>B</i>	0.7	-1.6	-0.2	-1.2	-0.1	3.4	3.7	0.8	0.0	2.0
Difference	1.2	2.4	1.3	1.3	0.0	1.0	1.8	0.8	4.6	1.4

Using the sign test, find out whether the difference is significant.

3. Assuming that the populations corresponding to the samples in Prob. 2 are normal, apply the test explained in Ex. 3, Sec. 18.15.
4. Complete the table in Ex. 2.
5. Make a table similar to that in Ex. 2 when $n = 4$.

6. Apply the test in Ex. 2 to the following data (from the Wörther Lake in Austria, x = depth in meters, y = oxygen content in milligrams per liter).

x	15	20	30	40	50	60	70
y	6.5	5.6	5.4	6.0	4.6	1.4	0.1

7. Apply the test in Ex. 2 to the following data (x = disulfide content of a certain type of wool, measured in per cent; y = maximum possible water content of the wool, measured in per cent).

x	10	15	30	40	50	55	80	100
y	50	46	43	42	36	39	37	33

18.20 Pairs of Measurements. Fitting Straight Lines

If in an experiment we measure two quantities X and Y simultaneously and the experiment is performed n times, then we obtain n pairs of values $(x_1, y_1), \dots, (x_n, y_n)$. In practice we distinguish between two types of experiments, as follows.

1. In *correlation analysis* both X and Y are random variables, and we are interested in relations between these variables. (We shall not discuss this branch of statistics.)

2. In *regression analysis* X is an ordinary variable (sometimes called a *controlled variable*), while Y depends on X and upon random influences so that Y is a random variable, and we are interested in the dependence of Y on X .

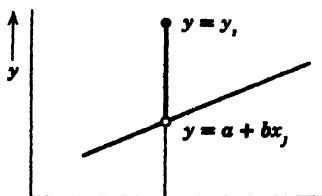


Fig. 419. Vertical distance of a point (x_i, y_i) from a line $y = a + bx$.

In the present section we shall consider a few basic facts from *linear regression analysis*, which arises if the dependence of Y on X is (nearly) linear. Then we would like to fit a straight line through the points representing $(x_1, y_1), \dots, (x_n, y_n)$ where x_1, \dots, x_n are values selected before the experiment is performed and y_1, \dots, y_n are corresponding observed values. This line may then be used for

estimating the mean value of Y corresponding to given values of X . If those points lie nearly on a straight line, then that line may be fitted "by eye." If they are more scattered, then we need a mathematical principle that yields a line of "best fit." Otherwise different people would probably draw quite different lines. Such a principle is the **method of least squares**, which we shall now introduce.

Consider a sample $(x_1, y_1), \dots, (x_n, y_n)$ of size n . The vertical distance (distance measured in y -direction) of a sample value (x_j, y_j) from a straight line $y = a + bx$ is $|y_j - a - bx_j|$; cf. Fig. 419. Hence the sum of the squares of these distances is

$$(1) \quad q = \sum_{j=1}^n (y_j - a - bx_j)^2.$$

In the method of least squares we choose a and b such that q is minimum. q depends on a and b , and a necessary condition for q to be minimum is

$$(2) \quad \frac{\partial q}{\partial a} = 0 \quad \text{and} \quad \frac{\partial q}{\partial b} = 0.$$

We shall see that from this condition we obtain

$$(3) \quad y - \bar{y} = b(x - \bar{x})$$

where

$$(4) \quad \bar{x} = \frac{1}{n}(x_1 + \dots + x_n) \quad \text{and} \quad \bar{y} = \frac{1}{n}(y_1 + \dots + y_n).$$

(3) is called the regression line of y on x . Its slope b is called the regression coefficient of y on x , and we shall see that

$$(5) \quad b = \frac{s_{xy}}{s_x^2}.$$

Here,

$$(6) \quad s_x^2 = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^2 = \frac{1}{n-1} \left[\sum_{j=1}^n x_j^2 - \frac{1}{n} \left(\sum_{j=1}^n x_j \right)^2 \right]$$

and

$$(7) \quad s_{xy} = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y}) = \frac{1}{n-1} \left[\sum_{j=1}^n x_j y_j - \frac{1}{n} \left(\sum_{j=1}^n x_j \right) \left(\sum_{j=1}^n y_j \right) \right].$$

s_{xy} is called the covariance of the sample. We see that the regression line (3) passes through the point (\bar{x}, \bar{y}) .

To derive (3), we use (1) and (2), finding

$$\frac{\partial q}{\partial a} = -2 \sum (y_j - a - bx_j) = 0$$

$$\frac{\partial q}{\partial b} = -2 \sum x_j (y_j - a - bx_j) = 0.$$

(where we sum over j from 1 to n). Thus

$$na + b \sum x_j = \sum y_j$$

$$a \sum x_j + b \sum x_j^2 = \sum x_j y_j.$$

We assume that the x -values are not all equal. Then the determinant

$$n \sum x_j^2 - (\sum x_j)^2 = n(n-1)s_x^2$$

[cf. (6)] of this system of linear equations is not zero, and the system has a unique solution [cf. (4), (6), (7)]

$$(8) \quad a = \bar{y} - b\bar{x}, \quad b = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n(n-1)s_1^2}.$$

This yields (3) with b given by (5)–(7).

The numerical work can be simplified by *coding*, that is, by setting

$$(9) \quad x_j = c_1 x_j^* + l_1, \quad y_j = c_2 y_j^* + l_2$$

and choosing the constant c_1, c_2, l_1, l_2 such that the transformed values x_j^* and y_j^* are as simple as possible. We first compute the values $\bar{x}^*, \bar{y}^*, s_{1^*}^2, s_{xy}^*$ corresponding to the transformed values and then

$$(10) \quad \bar{x} = c_1 \bar{x}^* + l_1, \quad \bar{y} = c_2 \bar{y}^* + l_2$$

$$s_1^2 = c_1^2 s_{1^*}^2, \quad s_{xy} = c_1 c_2 s_{xy}^*.$$

Example 1. The decrease of volume y [%] of leather for certain fixed values of high pressure x [atmospheres] was measured. The results are shown in the first two columns of Table 45. Find the regression line of y on x . We see that $n = 4$, and the first four columns of Table 45 yield

$$\bar{x} = \frac{28\ 000}{4} = 7\ 000, \quad \bar{y} = \frac{19.0}{4} = 4.75$$

$$s_{1^*}^2 = \frac{1}{3} \left(216\ 000\ 000 - \frac{28\ 000^2}{4} \right) = \frac{20\ 000\ 000}{3}$$

$$s_{xy} = \frac{1}{3} \left(148\ 400 - \frac{28\ 000 \cdot 19}{4} \right) = \frac{15\ 400}{3}.$$

Hence $b = 15\ 400 / 20\ 000\ 000 = 0.000\ 77$, and the regression line is

$$y - 4.75 = 0.000\ 77(x - 7000) \quad \text{or} \quad y = 0.000\ 77x - 0.64.$$

This computation becomes simpler if we set

$$x_j = 2000 x_j^* + 4000, \quad \text{thus} \quad x_j^* = 0.0005 x_j - 2$$

$$y_j = 0.1 y_j^* + 5, \quad \text{thus} \quad y_j^* = 10 y_j - 50.$$

Table 45. Regression of the Decrease of Volume y [%] of Leather on the Pressure x [Atmospheres].

Given values		Auxiliary values		Coded values		Auxiliary values	
x_j	y_j	x_j^2	$x_j y_j$	x_j^*	y_j^*	x_j^{*2}	$x_j^* y_j^*$
4 000	2.3	16 000 000	9 200	0	-27	0	0
6 000	4.1	36 000 000	24 600	1	-9	1	-9
8 000	5.7	64 000 000	45 600	2	7	4	14
10 000	6.9	100 000 000	69 000	3	19	9	57
Sums	28 000	19.0	216 000 000	148 400	6	-10	14
							62

(C. E. WEIR, Journ. Res. Nat. Bureau of Standards 45, 1950, 468).

From the last four columns of Table 45 we then obtain

$$\bar{x}^* = \frac{6}{4} = 1.5, \quad \bar{y}^* = -\frac{10}{4} = -2.5$$

$$s_{1^{*2}} = \frac{1}{3} \left(14 - \frac{6^2}{4} \right) = \frac{5}{3}, \quad s_{xy}^* = \frac{1}{3} \left(62 - \frac{6 \cdot (-10)}{4} \right) = \frac{77}{3}$$

Using (10) with $c_1 = 2000$, $c_2 = 0.1$, $l_1 = 4000$, $l_2 = 5$, we find

$$\bar{x} = 2000 \cdot 1.5 + 4000 = 7000, \quad \bar{y} = 0.1 \cdot (-2.5) + 5 = 4.75$$

$$b = \frac{s_{xy}}{s_{1^{*2}}} = \frac{c_2 s_{xy}^*}{c_1 s_{1^{*2}}} = \frac{0.1 \cdot 77/3}{2000 \cdot 5/3} = 0.00077,$$

as before.

We remember that in our present consideration X is an ordinary variable (controlled variable) while Y is a random variable. We now make the following

Assumption. For any given value $X = x$ the random variable Y is normal with mean

$$(11) \quad \mu = \alpha + \beta x$$

and variance σ^2 , where α , β , σ^2 are parameters which do not depend on x .

Table 46. Determination of a Confidence Interval for β in (11) under the Preceding Assumption.

1st step. Choose a confidence level γ (95%, 99% or the like).

2nd step. Determine the solution c of the equation

$$(12) \quad F(c) = \frac{1}{2}(1 + \gamma)$$

from the table of the t -distribution with $n - 2$ degrees of freedom (Table 52, Sec. 18.21; n = sample size).

3rd step. Using a sample $(x_1, y_1), \dots, (x_n, y_n)$, compute s_1^2 [cf. (6)], b [cf. (5), (7)],

$$(13) \quad s_2^2 = \frac{1}{n-1} \left[\sum_{j=1}^n y_j^2 - \frac{1}{n} \left(\sum_{j=1}^n y_j \right)^2 \right]$$

and

$$(14) \quad q_0 = (n-1)(s_2^2 - b^2 s_1^2).$$

4th step. Compute $l = c\sqrt{q_0/s_1^2} \sqrt{(n-1)(n-2)}$. The confidence interval is

$$(15) \quad \text{CONF } \{b - l \leq \beta \leq b + l\}.$$

Then we may use a sample for obtaining a confidence interval for β , as shown in Table 46. To justify this procedure, we may regard y_i in (x_i, y_i) as an observed value of a normal random variable Y_i with mean $\alpha + \beta x_i$ and variance σ^2 . Since the sample values are independent, those variables Y_1, \dots, Y_n are independent, too. We may now regard b as an observed value of the random variable

$$B = \frac{1}{(n-1)s_1^2} \sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y}) \quad \left[\bar{Y} = \frac{1}{n}(Y_1 + \dots + Y_n) \right].$$

s_2^2 as an observed value of the random variable

$$S_2^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2 = \frac{1}{n-1} \left[\sum_{i=1}^n Y_i^2 - \frac{1}{n} \left(\sum_{i=1}^n Y_i \right)^2 \right].$$

and Q_0 as an observed value of the random variable

$$Q_0 = (n-1)(S_2^2 - B^2 s_1^2) \quad [\text{cf. (6)}].$$

It can be shown that under the preceding assumption the random variable

$$T = s_1 \sqrt{(n-1)(n-2)} \frac{B - \beta}{\sqrt{Q_0}}$$

has a t -distribution with $n-2$ degrees of freedom (proof in Ref. [13], cf. Appendix 1). From

$$P(-c \leq T \leq c) = F(c) - F(-c) = \gamma$$

we obtain (12), because the t -distribution is symmetric. If we insert the expression for T into $-c \leq T \leq c$, transform this into an inequality for β , and replace B and Q_0 by those observed values, (15) follows.

Note that (14) is identical with (1) referred to the regression line. In fact, by inserting a [cf. (8)] into (1) we obtain

$$\begin{aligned} q &= \sum [(y_i - \bar{y}) - b(x_i - \bar{x})]^2 \\ &= \sum (y_i - \bar{y})^2 - 2b \sum (y_i - \bar{y})(x_i - \bar{x}) + b^2 \sum (x_i - \bar{x})^2 \\ &= (n-1)s_2^2 - 2b(n-1)s_{xy} + b^2(n-1)s_1^2. \end{aligned}$$

From (5) we see that $s_{xy} = bs_1^2$, and the expression on the right reduces to the right-hand side of (14).

Example 2. Using the sample in Table 45, determine a confidence interval for β by the method in Table 46.

1st step. We choose $\gamma = 0.95$.

2nd step. Equation (12) takes the form $F(c) = 0.975$, and Table 52 with $n - 2 = 2$ degrees of freedom yields $c = 4.30$.

3rd step. From Ex. 1 we have $s_1^2 = 20\,000\,000/3$ and $b = 0.000\,77$. From Table 45 we compute

$$s_2^2 = \frac{1}{3} \left(102.2 - \frac{19^2}{4} \right) = \frac{11.95}{3}$$

and

$$q_0 = 3s_2^2 - 3s_1^2 b^2 = 11.95 - 20\,000\,000 \cdot 0.000\,77^2 = 0.092.$$

4th step. We thus obtain

$$l = 4.30 \cdot 0.3033 / \sqrt{2 \cdot 20\,000\,000} = 0.000\,206.$$

Hence the confidence interval is

$$\text{CONF } \{0.000\,56 \leq \beta \leq 0.000\,98\}.$$

Problems

In each case find and graph the regression line of y on x .

1. The sample in Prob. 6, Sec. 18.19
2. The sample in Prob. 7, Sec. 18.19
3. Number of errors per hour of radio telegraphists as a function of the temperature x [$^{\circ}\text{C}$]

x	26	28	31	33
y	12.0	11.5	15.3	17.3

4. Mortality y [%] of termites (*Reticulitermes lucifugus Rossi*) as a function of the concentration x [%] of chloronaphthalene

x	0.04	0.15	0.30	1.00	2.00
y	3	16	13	70	90

5. The iron content y [%] of a certain type of iron ore was measured as a function of the density x [grams/cm³]. The results were as follows.

x	2.8	2.9	3.0	3.1	3.2	3.2	3.2	3.3	3.4
y	27	23	30	28	30	32	34	33	30

Suppose that the corresponding population satisfies the assumptions used in Table 46 and determine a 95%-confidence interval for β .

6. Find a 99%-confidence interval for β using the sample in Prob. 4. (Suppose that the corresponding population satisfies the assumption used in Table 46.)

18.21 Statistical Tables

Table 47. Binomial Distribution

Probability function $f(x)$ (cf. (2); Sec. 18.9) and distribution function $F(x)$

n	x	$p = 0.1$		$p = 0.2$		$p = 0.3$		$p = 0.4$		$p = 0.5$	
		$f(x)$	$F(x)$								
1	0	0.		0.		0.		0.		0.	
	1	9000	0.9000	8000	0.8000	7000	0.7000	6000	0.6000	5000	0.5000
2	0	8100	0.8100	6400	0.6400	4900	0.4900	3600	0.3600	2500	0.2500
	1	1800	0.9900	3200	0.9600	4200	0.9100	4800	0.8400	5000	0.7500
3	2	0100	1.0000	0400	1.0000	0900	1.0000	1600	1.0000	2500	1.0000
	0	7290	0.7290	5120	0.5120	3430	0.3430	2160	0.2160	1250	0.1250
4	1	2430	0.9720	3840	0.8960	4410	0.7840	4320	0.6480	3750	0.5000
	2	0270	0.9990	0960	0.9920	1890	0.9730	2880	0.9360	3750	0.8750
5	3	0010	1.0000	0080	1.0000	0270	1.0000	0640	1.0000	1250	1.0000
	0	6561	0.6561	4096	0.4096	2401	0.2401	1296	0.1296	625	0.0625
6	1	2916	0.9477	4096	0.8192	4116	0.6517	3456	0.4752	2500	0.3125
	2	0486	0.9963	1536	0.9728	2646	0.9163	3456	0.8208	3750	0.6875
7	3	0036	0.9999	0256	0.9984	0756	0.9919	1536	0.9744	2500	0.9375
	4	0001	1.0000	0016	1.0000	0081	1.0000	0256	1.0000	625	1.0000
5	0	5905	0.5905	3277	0.3277	1681	0.1681	0778	0.0778	0313	0.0313
	1	3281	0.9185	4096	0.7373	3602	0.5282	2592	0.3370	1563	0.1875
6	2	0729	0.9914	2048	0.9421	3087	0.8369	3456	0.6826	3125	0.5000
	3	0081	0.9995	0512	0.9933	1323	0.9692	2304	0.9130	3125	0.8125
7	4	0005	1.0000	0064	0.9997	0284	0.9976	0768	0.9898	1563	0.9688
	5	0000	1.0000	0003	1.0000	0024	1.0000	0102	1.0000	0313	1.0000
6	0	5314	0.5314	2621	0.2621	1176	0.1176	0467	0.0467	0156	0.0156
	1	3543	0.8857	3932	0.6554	3025	0.4202	1866	0.2333	0938	0.1094
7	2	0984	0.9841	2458	0.9011	3241	0.7443	3110	0.5443	2344	0.3438
	3	0146	0.9987	0819	0.9830	1852	0.9295	2765	0.8208	3125	0.6563
8	4	0012	0.9999	0154	0.9984	0595	0.9891	1382	0.9590	2344	0.8906
	5	0001	1.0000	0015	0.9999	0102	0.9993	0369	0.9959	0938	0.9844
9	6	0000	1.0000	0001	1.0000	0007	1.0000	0041	1.0000	0156	1.0000
	7	0000	1.0000	0000	1.0000	0002	1.0000	0016	1.0000	0078	1.0000
7	0	4783	0.4783	2097	0.2097	0824	0.0824	0280	0.0280	0078	0.0078
	1	3720	0.8503	3670	0.5767	2471	0.3294	1306	0.1586	0547	0.0625
8	2	1240	0.9743	2753	0.8520	3177	0.6471	2613	0.4199	1641	0.2266
	3	0230	0.9973	1147	0.9667	2269	0.8740	2903	0.7102	2734	0.5000
9	4	0026	0.9998	0287	0.9953	0972	0.9712	1935	0.9037	2734	0.7734
	5	0002	1.0000	0043	0.9996	0250	0.9962	0774	0.9812	1641	0.9375
10	6	0000	1.0000	0004	1.0000	0036	0.9998	0172	0.9984	0547	0.9922
	7	0000	1.0000	0000	1.0000	0002	1.0000	0016	1.0000	0078	1.0000
8	0	4305	0.4305	1678	0.1678	0576	0.0576	0168	0.0168	0039	0.0039
	1	3826	0.8131	3355	0.5033	1977	0.2553	0896	0.1064	0313	0.0352
9	2	1488	0.9619	2936	0.7969	2965	0.5518	2090	0.3154	1094	0.1445
	3	0331	0.9950	1468	0.9437	2541	0.8059	2787	0.5941	2188	0.3633
10	4	0046	0.9996	0459	0.9896	1361	0.9420	2322	0.8263	2734	0.6367
	5	0004	1.0000	0092	0.9988	0467	0.9887	1239	0.9502	2188	0.8555
11	6	0000	1.0000	0011	0.9999	0100	0.9987	0413	0.9915	1094	0.9648
	7	0000	1.0000	0001	1.0000	0012	0.9999	0079	0.9993	0313	0.9961
12	8	0000	1.0000	0000	1.0000	0001	1.0000	0007	1.0000	0039	1.0000

Table 48. Poisson DistributionProbability function $f(x)$ (cf. (5), Sec. 18.9) and distribution function $F(x)$

x	$\mu = 0.1$		$\mu = 0.2$		$\mu = 0.3$		$\mu = 0.4$		$\mu = 0.5$	
	$f(x)$	$F(x)$								
0	0.9048	0.9048	0.8187	0.8187	0.7408	0.7408	0.6703	0.6703	0.6065	0.6065
1	0.0905	0.9953	0.1637	0.9825	0.2222	0.9631	0.2681	0.9384	0.3033	0.9089
2	0.0045	0.9998	0.0164	0.9989	0.0333	0.9964	0.0536	0.9921	0.0758	0.9856
3	0.0002	1.0000	0.0011	0.9999	0.0033	0.9997	0.0072	0.9992	0.0126	0.9982
4	0.0000	1.0000	0.0001	1.0000	0.0003	1.0000	0.0007	0.9999	0.0016	0.9998
5							0.0001	1.0000	0.0002	1.0000

x	$\mu = 0.6$		$\mu = 0.7$		$\mu = 0.8$		$\mu = 0.9$		$\mu = 1$	
	$f(x)$	$F(x)$	$f(x)$	$F(x)$	$f(x)$	$F(x)$	$f(x)$	$F(x)$	$f(x)$	$F(x)$
0	0.5488	0.5488	0.4966	0.4966	0.4493	0.4493	0.4066	0.4066	0.3679	0.3679
1	3293	0.8781	3476	0.8442	3595	0.8088	3659	0.7725	3679	0.7358
2	0988	0.9769	1217	0.9659	1438	0.9526	1647	0.9371	1839	0.9197
3	0198	0.9966	0284	0.9942	0383	0.9909	0494	0.9865	0613	0.9810
4	0030	0.9996	0050	0.9992	0077	0.9986	0111	0.9977	0153	0.9963
5	0004	1.0000	0007	0.9999	0012	0.9998	0020	0.9997	0031	0.9994
6			0001	1.0000	0002	1.0000	0003	1.0000	0005	0.9999
7							0001	1.0000		

x	$\mu = 1.5$		$\mu = 2$		$\mu = 3$		$\mu = 4$		$\mu = 5$	
	$f(x)$	$F(x)$	$f(x)$	$F(x)$	$f(x)$	$F(x)$	$f(x)$	$F(x)$	$f(x)$	$F(x)$
0	0.2231	0.2231	0.1353	0.1353	0.0498	0.0498	0.0183	0.0183	0.0067	0.0067
1	3347	0.5578	2707	0.4060	1494	0.1991	0733	0.0916	0337	0.0404
2	2510	0.8088	2707	0.6767	2240	0.4232	1465	0.2381	0842	0.1247
3	1255	0.9344	1804	0.8571	2240	0.6472	1954	0.4335	1404	0.2650
4	0471	0.9814	0902	0.9473	1680	0.8153	1954	0.6288	1755	0.4405
5	0141	0.9955	0361	0.9834	1008	0.9161	1563	0.7851	1755	0.6160
6	0035	0.9991	0120	0.9955	0504	0.9665	1042	0.8893	1462	0.7622
7	0008	0.9998	0034	0.9989	0216	0.9881	0595	0.9489	1044	0.8666
8	0001	1.0000	0009	0.9998	0081	0.9962	0298	0.9786	0653	0.9319
9			0002	1.0000	0027	0.9989	0132	0.9919	0363	0.9682
10					0008	0.9997	0053	0.9972	0181	0.9863
11					0002	0.9999	0019	0.9991	0082	0.9945
12					0001	1.0000	0006	0.9997	0034	0.9980
13							0002	0.9999	0013	0.9993
14							0001	1.0000	0005	0.9998
15									0002	0.9999
16									0000	1.0000

Table 49. Normal DistributionValues of the distribution function $\Phi(z)$ (cf. (4), Sec. 18.10) $\Phi(-z) = 1 - \Phi(z)$, $\Phi(0) = 0.5000$

z	$\Phi(z)$										
0.01	0.5040	0.51	0.6950	1.01	0.8438	1.51	0.9345	2.01	0.9778	2.51	0.9940
0.02	0.5080	0.52	0.6985	1.02	0.8461	1.52	0.9357	2.02	0.9783	2.52	0.9941
0.03	0.5120	0.53	0.7019	1.03	0.8485	1.53	0.9370	2.03	0.9788	2.53	0.9943
0.04	0.5160	0.54	0.7054	1.04	0.8508	1.54	0.9382	2.04	0.9793	2.54	0.9945
0.05	0.5199	0.55	0.7088	1.05	0.8531	1.55	0.9394	2.05	0.9798	2.55	0.9946
0.06	0.5239	0.56	0.7123	1.06	0.8554	1.56	0.9406	2.06	0.9803	2.56	0.9948
0.07	0.5279	0.57	0.7157	1.07	0.8577	1.57	0.9418	2.07	0.9808	2.57	0.9949
0.08	0.5319	0.58	0.7190	1.08	0.8599	1.58	0.9429	2.08	0.9812	2.58	0.9951
0.09	0.5359	0.59	0.7224	1.09	0.8621	1.59	0.9441	2.09	0.9817	2.59	0.9952
0.10	0.5398	0.60	0.7257	1.10	0.8643	1.60	0.9452	2.10	0.9821	2.60	0.9953
0.11	0.5438	0.61	0.7291	1.11	0.8665	1.61	0.9463	2.11	0.9826	2.61	0.9955
0.12	0.5478	0.62	0.7324	1.12	0.8686	1.62	0.9474	2.12	0.9830	2.62	0.9956
0.13	0.5517	0.63	0.7357	1.13	0.8708	1.63	0.9484	2.13	0.9834	2.63	0.9957
0.14	0.5557	0.64	0.7389	1.14	0.8729	1.64	0.9495	2.14	0.9838	2.64	0.9959
0.15	0.5596	0.65	0.7422	1.15	0.8749	1.65	0.9505	2.15	0.9842	2.65	0.9960
0.16	0.5636	0.66	0.7454	1.16	0.8770	1.66	0.9515	2.16	0.9846	2.66	0.9961
0.17	0.5675	0.67	0.7486	1.17	0.8790	1.67	0.9525	2.17	0.9850	2.67	0.9962
0.18	0.5714	0.68	0.7517	1.18	0.8810	1.68	0.9535	2.18	0.9854	2.68	0.9963
0.19	0.5753	0.69	0.7549	1.19	0.8830	1.69	0.9545	2.19	0.9857	2.69	0.9964
0.20	0.5793	0.70	0.7580	1.20	0.8849	1.70	0.9554	2.20	0.9861	2.70	0.9965
0.21	0.5832	0.71	0.7611	1.21	0.8869	1.71	0.9564	2.21	0.9864	2.71	0.9966
0.22	0.5871	0.72	0.7642	1.22	0.8888	1.72	0.9573	2.22	0.9868	2.72	0.9967
0.23	0.5910	0.73	0.7673	1.23	0.8907	1.73	0.9582	2.23	0.9871	2.73	0.9968
0.24	0.5948	0.74	0.7704	1.24	0.8925	1.74	0.9591	2.24	0.9875	2.74	0.9969
0.25	0.5987	0.75	0.7734	1.25	0.8944	1.75	0.9599	2.25	0.9878	2.75	0.9970
0.26	0.6026	0.76	0.7764	1.26	0.8962	1.76	0.9608	2.26	0.9881	2.76	0.9971
0.27	0.6064	0.77	0.7794	1.27	0.8980	1.77	0.9616	2.27	0.9884	2.77	0.9972
0.28	0.6103	0.78	0.7823	1.28	0.8997	1.78	0.9625	2.28	0.9887	2.78	0.9973
0.29	0.6141	0.79	0.7852	1.29	0.9015	1.79	0.9633	2.29	0.9890	2.79	0.9974
0.30	0.6179	0.80	0.7881	1.30	0.9032	1.80	0.9641	2.30	0.9893	2.80	0.9974
0.31	0.6217	0.81	0.7910	1.31	0.9049	1.81	0.9649	2.31	0.9896	2.81	0.9975
0.32	0.6255	0.82	0.7939	1.32	0.9066	1.82	0.9656	2.32	0.9898	2.82	0.9976
0.33	0.6293	0.83	0.7967	1.33	0.9082	1.83	0.9664	2.33	0.9901	2.83	0.9977
0.34	0.6331	0.84	0.7995	1.34	0.9099	1.84	0.9671	2.34	0.9904	2.84	0.9977
0.35	0.6368	0.85	0.8023	1.35	0.9115	1.85	0.9678	2.35	0.9906	2.85	0.9978
0.36	0.6406	0.86	0.8051	1.36	0.9131	1.86	0.9686	2.36	0.9909	2.86	0.9979
0.37	0.6443	0.87	0.8078	1.37	0.9147	1.87	0.9693	2.37	0.9911	2.87	0.9979
0.38	0.6480	0.88	0.8106	1.38	0.9162	1.88	0.9699	2.38	0.9913	2.88	0.9980
0.39	0.6517	0.89	0.8133	1.39	0.9177	1.89	0.9706	2.39	0.9916	2.89	0.9981
0.40	0.6554	0.90	0.8159	1.40	0.9192	1.90	0.9713	2.40	0.9918	2.90	0.9981
0.41	0.6591	0.91	0.8186	1.41	0.9207	1.91	0.9719	2.41	0.9920	2.91	0.9982
0.42	0.6628	0.92	0.8212	1.42	0.9222	1.92	0.9726	2.42	0.9922	2.92	0.9982
0.43	0.6664	0.93	0.8238	1.43	0.9236	1.93	0.9732	2.43	0.9925	2.93	0.9983
0.44	0.6700	0.94	0.8264	1.44	0.9251	1.94	0.9738	2.44	0.9927	2.94	0.9984
0.45	0.6736	0.95	0.8289	1.45	0.9265	1.95	0.9744	2.45	0.9929	2.95	0.9984
0.46	0.6772	0.96	0.8315	1.46	0.9279	1.96	0.9750	2.46	0.9931	2.96	0.9985
0.47	0.6808	0.97	0.8340	1.47	0.9292	1.97	0.9756	2.47	0.9932	2.97	0.9985
0.48	0.6844	0.98	0.8365	1.48	0.9306	1.98	0.9761	2.48	0.9934	2.98	0.9986
0.49	0.6879	0.99	0.8389	1.49	0.9319	1.99	0.9767	2.49	0.9936	2.99	0.9986
0.50	0.6915	1.00	0.8413	1.50	0.9332	2.00	0.9772	2.50	0.9938	3.00	0.9987

Table 50. Normal Distribution

Values of z for given values of $\Phi(z)$ (cf. (4), Sec. 18.10) and $D(z) = \Phi(z) - \Phi(-z)$ Example: $z = 0.279$ when $\Phi(z) = 61\%$; $z = 0.860$ when $D(z) = 61\%$.

%	$z(\Phi)$	$z(D)$	%	$z(\Phi)$	$z(D)$	%	$z(\Phi)$	$z(D)$
1	-2.326	0.013	41	-0.228	0.539	81	0.878	1.311
2	-2.054	0.025	42	-0.202	0.553	82	0.915	1.341
3	-1.881	0.038	43	-0.176	0.568	83	0.954	1.372
4	-1.751	0.050	44	-0.151	0.583	84	0.994	1.405
5	-1.645	0.063	45	-0.126	0.598	85	1.036	1.440
6	-1.555	0.075	46	-0.100	0.613	86	1.080	1.476
7	-1.476	0.088	47	-0.075	0.628	87	1.126	1.514
8	-1.405	0.100	48	-0.050	0.643	88	1.175	1.555
9	-1.341	0.113	49	-0.025	0.659	89	1.227	1.598
10	-1.282	0.126	50	0.000	0.674	90	1.282	1.645
11	-1.227	0.138	51	0.025	0.690	91	1.341	1.695
12	-1.175	0.151	52	0.050	0.706	92	1.405	1.751
13	-1.126	0.164	53	0.075	0.722	93	1.476	1.812
14	-1.080	0.176	54	0.100	0.739	94	1.555	1.881
15	-1.036	0.189	55	0.126	0.755	95	1.645	1.960
16	-0.994	0.202	56	0.151	0.772	96	1.751	2.054
17	-0.954	0.215	57	0.176	0.789	97	1.881	2.170
18	-0.915	0.228	58	0.202	0.806	98	2.054	2.326
19	-0.878	0.240	59	0.228	0.824	99	2.326	2.576
20	-0.842	0.253	60	0.253	0.842			
21	-0.806	0.266	61	0.279	0.860	99.1	2.366	2.612
22	-0.772	0.279	62	0.305	0.878	99.2	2.409	2.652
23	-0.739	0.292	63	0.332	0.896	99.3	2.457	2.697
24	-0.706	0.305	64	0.358	0.915	99.4	2.512	2.748
25	-0.674	0.319	65	0.385	0.935	99.5	2.576	2.807
26	-0.643	0.332	66	0.412	0.954	99.6	2.652	2.878
27	-0.613	0.345	67	0.440	0.974	99.7	2.748	2.968
28	-0.583	0.358	68	0.468	0.994	99.8	2.878	3.090
29	-0.553	0.372	69	0.496	1.015	99.9	3.090	3.291
30	-0.524	0.385	70	0.524	1.036			
31	-0.496	0.399	71	0.553	1.058	99.91	3.121	3.320
32	-0.468	0.412	72	0.583	1.080	99.92	3.156	3.353
33	-0.440	0.426	73	0.613	1.103	99.93	3.195	3.390
34	-0.412	0.440	74	0.643	1.126	99.94	3.239	3.432
35	-0.385	0.454	75	0.674	1.150	99.95	3.291	3.481
36	-0.358	0.468	76	0.706	1.175	99.96	3.353	3.540
37	-0.332	0.482	77	0.739	1.200	99.97	3.432	3.615
38	-0.305	0.496	78	0.772	1.227	99.98	3.540	3.719
39	-0.279	0.510	79	0.806	1.254	99.99	3.719	3.891
40	-0.253	0.524	80	0.842	1.282			

Table 51. Random Digits

Cf. Sec. 18.12

Row No.	Column Number									
	0	1	2	3	4	5	6	7	8	9
0	87331	82442	28104	26432	83640	17323	68764	84728	37995	96106
1	33628	17364	01409	87803	65641	33433	48944	64299	79066	31777
2	54680	13427	72496	16967	16195	96593	55040	53729	62035	66717
3	51199	49794	49407	10774	98140	83891	37195	24066	61140	65144
4	78702	98067	61313	91661	59861	54437	77739	19892	54817	88645
5	55672	16014	24892	13089	00410	81458	76156	28189	40595	21500
6	18880	58497	03862	32368	59320	24807	63392	79793	63043	09425
7	10242	62548	62330	05703	33535	49128	66298	16193	55301	01306
8	54993	17182	94618	23228	83895	73251	68199	64639	83178	70521
9	22686	50885	16006	04041	08077	33065	35237	02502	94755	72062
10	42349	03145	15770	70665	53291	32288	41568	66079	98705	31029
11	18093	09553	39428	75464	71329	86344	80729	40916	18860	51780
12	11535	03924	84252	74795	40193	84597	42497	21918	91384	84721
13	35066	73848	65351	53270	67341	70177	92373	17604	42204	60476
14	57477	22809	73558	96182	96779	01604	25748	59553	64876	94611
15	48647	33850	52956	45410	88212	05120	99391	32276	55961	41775
16	86857	81154	22223	74950	53296	67767	55866	49061	66937	81818
17	20182	36907	94644	99122	09774	29189	27212	79000	50217	71077
18	83687	31231	01133	41432	54542	60204	81618	09586	34481	87683
19	81315	12390	46074	47810	90171	36313	95440	77583	28506	38808
20	87026	52826	58341	76549	04105	66191	12914	55348	07907	06978
21	34301	76733	07251	90524	21931	83695	41340	53581	64582	60210
22	70734	24337	32674	49508	49751	90489	63202	24380	77943	09942
23	94710	31527	73445	32839	68176	53580	85250	53243	03350	00128
24	76462	16987	07775	43162	11777	16810	75158	13894	88945	15539
25	14348	28403	79245	69023	34196	46398	05964	64715	11330	17515
26	74618	89317	30146	25606	94507	98104	04239	44973	37636	88866
27	99442	19200	85406	45358	86253	60638	38858	44964	54103	57287
28	26869	44399	89452	06652	31271	00647	46551	83050	92058	83814
29	80988	08149	50499	98584	28385	63680	44638	91864	96002	87802
30	07511	79047	89289	17774	67194	37362	85684	55505	97809	67056
31	49779	12138	05048	03535	27502	63308	10218	53296	48687	61340
32	47938	55945	24003	19635	17471	65997	85906	98694	56420	78357
33	15604	06626	14360	79542	13512	87595	08542	01800	35443	52823
34	12307	27726	21864	00045	16075	03770	86978	52718	02693	09096
35	02450	28053	66134	99445	91316	25727	89399	85272	67148	78358
36	57623	54382	35236	89244	27245	90500	75430	96762	71968	65838
37	91762	78849	93105	40481	99431	03304	21079	86459	21287	76566
38	87373	31137	31128	67050	34309	44914	80711	61738	61498	24288
39	67094	41485	54149	86088	10192	21174	39948	67268	29938	32476
40	94456	66747	76922	87627	71834	57688	04878	78348	68970	60048
41	68359	75292	27710	86889	81678	79798	58360	39175	75667	65782
42	52393	31404	32584	06837	79762	13168	76055	54833	22841	98889
43	59565	91254	11847	20672	37625	41454	86861	55824	79793	74575
44	48185	11066	20162	38230	16043	48409	47421	21195	98008	57305
45	19230	12187	86659	12971	52204	76546	63272	19312	81667	96557
46	84327	21942	81727	68735	89190	58491	55329	96875	19463	89687
47	77430	71210	00591	50124	12030	50280	12358	76174	48353	09682
48	12462	19108	70512	53926	25595	97085	03833	59806	12351	64253
49	11684	06644	57816	10078	45021	47751	38285	73520	08434	65627,

Table 51. Random Digits (continued)

Row No.	Column Number									
	0	1	2	3	4	5	6	7	8	9
50	12896	36576	68686	08462	65652	76571	70891	09007	04581	01684
51	59090	05111	27587	90349	30789	50304	70650	06646	70126	15284
52	424286	67483	65282	19037	80588	73076	41820	46651	40442	40718
53	88662	03928	03249	85910	97533	88643	29829	21557	47328	36724
54	69403	03626	92678	53460	15465	83516	54012	80509	55976	46115
55	56434	70543	38696	98502	32092	95505	62091	39549	30117	98209
56	58227	62694	42837	29183	11393	68463	25150	86338	95620	39836
57	41272	94927	15413	40505	33123	63218	72940	98349	57249	40170
58	36819	01162	30425	15546	16065	68459	35776	64276	92868	07372
59	31700	66711	26115	55755	33584	18091	38709	57276	74660	90392
60	69855	63699	36839	90531	97125	87875	62824	03889	12538	24740
61	44322	17569	45439	41455	34324	90902	07978	26268	04279	76816
62	62226	36661	87011	66267	78777	78044	40819	49496	39814	73867
63	27284	19737	98741	72531	52741	26699	98755	19657	08665	16818
64	88341	21652	94743	77268	79525	44769	66583	30621	90534	62050
65	53266	18783	51903	56711	38060	69513	61963	80470	88018	86510
66	50527	49330	24839	42529	03944	95219	88724	37247	84166	23023
67	15655	07852	77206	35944	71446	30573	19405	57824	23576	23301
68	62057	22206	03314	83465	57466	10465	19891	32308	01900	67484
69	41769	56091	19892	96253	92808	45785	52774	49674	68103	65032
70	25993	72416	44473	41299	93095	17338	69802	98548	02429	85238
71	22842	57871	04470	37373	34516	04042	04078	35336	34393	97573
72	55704	31982	05234	22664	22181	40358	28089	15790	33340	18852
73	94258	18706	09437	96041	90052	80862	20420	24323	11635	91677
74	74145	20453	29657	98868	56695	53483	87449	35060	98942	62697
75	88881	12673	73961	89884	73247	97670	69570	88888	58560	72580
76	01508	56780	52223	35632	73347	71317	46541	88023	36656	76332
77	92069	43000	23233	06058	82527	25250	27555	20426	60361	63525
78	53366	35249	02117	68620	39388	69795	73215	01846	16983	78560
79	88057	54097	49511	74867	32192	90071	04147	46094	63519	07199
80	85492	82238	02668	91854	86149	28590	77853	81035	45561	16032
81	39453	62123	69611	53017	34964	09786	24614	49514	01056	18700
82	82627	98111	93870	56969	69566	62662	07353	84838	14570	14508
83	61142	51743	38209	31474	96095	15163	54380	77849	20465	03142
84	12031	32528	61311	53730	89032	16124	58844	35386	45521	59368
85	31313	59838	29147	76882	74328	09955	63673	96651	53264	29871
86	50767	41056	97409	44376	62219	35439	70102	99248	71179	26052
87	30522	95699	84966	26554	24768	72247	84993	85375	92518	16334
88	74176	19870	89874	64799	03792	57006	57225	36677	46825	14087
89	17114	93248	37065	91346	04657	93763	92210	43676	44944	75798
90	53005	11825	64608	87587	05742	31914	55044	41818	29667	77424
91	31985	81539	79942	49471	46200	27639	94099	42085	79231	03932
92	63499	60508	77522	15624	15088	78519	52279	79214	43623	69166
93	30506	42444	99047	66010	91657	37160	37408	85714	21420	80996
94	78248	16841	92357	10130	68990	38307	61022	56806	81016	38511
95	64996	84789	50185	32200	64382	29752	11876	00664	54547	62597
96	11963	13157	09136	01769	30117	71486	80111	09161	08371	71749
97	44335	91450	43456	90449	18338	19787	31339	60473	06606	89788
98	42277	11868	44520	01113	11341	11743	97949	49718	99176	42006
99	77562	18863	58515	90166	78508	14864	19111	57183	85808	59385

Table 52. t-Distribution

Values of z for given values of the distribution function $F(z)$ (cf. (4), Sec. 18.14). Example: For 9 degrees of freedom, $z = 1.83$ when $F(z) = 0.95$.

$F(z)$	Number of degrees of freedom									
	1	2	3	4	5	6	7	8	9	10
0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.6	0.33	0.29	0.28	0.27	0.27	0.27	0.26	0.26	0.26	0.26
0.7	0.73	0.62	0.58	0.57	0.56	0.55	0.55	0.55	0.54	0.54
0.8	1.38	1.06	0.98	0.94	0.92	0.91	0.90	0.89	0.88	0.88
0.9	3.08	1.89	1.64	1.53	1.48	1.44	1.42	1.40	1.38	1.37
0.95	6.31	2.92	2.35	2.13	2.02	1.94	1.90	1.86	1.83	1.81
0.975	12.7	4.30	3.18	2.78	2.57	2.45	2.37	2.31	2.26	2.23
0.99	31.8	6.97	4.54	3.75	3.37	3.14	3.00	2.90	2.82	2.76
0.995	63.7	9.93	5.84	4.60	4.03	3.71	3.50	3.36	3.25	3.17
0.999	318.3	22.3	10.2	7.17	5.89	5.21	4.79	4.50	4.30	4.14

$F(z)$	Number of degrees of freedom									
	11	12	13	14	15	16	17	18	19	20
0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.6	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26
0.7	0.54	0.54	0.54	0.54	0.54	0.54	0.53	0.53	0.53	0.53
0.8	0.88	0.87	0.87	0.87	0.87	0.87	0.86	0.86	0.86	0.86
0.9	1.36	1.36	1.35	1.35	1.34	1.34	1.33	1.33	1.33	1.33
0.95	1.80	1.78	1.77	1.76	1.75	1.75	1.74	1.73	1.73	1.73
0.975	2.20	2.18	2.16	2.15	2.13	2.12	2.11	2.10	2.09	2.09
0.99	2.72	2.68	2.65	2.62	2.60	2.58	2.57	2.55	2.54	2.53
0.995	3.11	3.06	3.01	2.98	2.95	2.92	2.90	2.88	2.86	2.85
0.999	4.03	3.93	3.85	3.79	3.73	3.69	3.65	3.61	3.58	3.55

$F(z)$	Number of degrees of freedom									
	22	24	26	28	30	40	50	100	200	∞
0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.6	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.25	0.25	0.25
0.7	0.53	0.53	0.53	0.53	0.53	0.53	0.53	0.53	0.53	0.52
0.8	0.86	0.86	0.86	0.86	0.85	0.85	0.85	0.85	0.84	0.84
0.9	1.32	1.32	1.32	1.31	1.31	1.30	1.30	1.29	1.29	1.28
0.95	1.72	1.71	1.71	1.70	1.70	1.68	1.68	1.66	1.65	1.65
0.975	2.07	2.06	2.06	2.05	2.04	2.02	2.01	1.98	1.97	1.96
0.99	2.51	2.49	2.48	2.47	2.46	2.42	2.40	2.37	2.35	2.33
0.995	2.82	2.80	2.78	2.76	2.75	2.70	2.68	2.63	2.60	2.58
0.999	3.51	3.47	3.44	3.41	3.39	3.31	3.26	3.17	3.13	3.09

Table 53. Chi-square Distribution

Values of z for given values of the distribution function $F(z)$ (cf. (7), Sec. 18.14). Example: For 3 degrees of freedom, $z = 11.34$ when $F(z) = 0.99$.

$F(z)$	Number of degrees of freedom									
	1	2	3	4	5	6	7	8	9	10
0.005	0.00	0.01	0.07	0.21	0.41	0.68	0.99	1.34	1.73	2.16
0.01	0.00	0.02	0.11	0.30	0.55	0.87	1.24	1.65	2.09	2.56
0.025	0.00	0.05	0.22	0.48	0.83	1.24	1.69	2.18	2.70	3.25
0.05	0.00	0.10	0.35	0.71	1.15	1.64	2.17	2.73	3.33	3.94
0.95	3.84	5.99	7.81	9.49	11.07	12.59	14.07	15.51	16.92	18.31
0.975	5.02	7.38	9.35	11.14	12.83	14.45	16.01	17.53	19.02	20.48
0.99	6.63	9.21	11.34	13.28	15.09	16.81	18.48	20.09	21.67	23.21
0.995	7.88	10.60	12.84	14.86	16.75	18.55	20.28	21.96	23.59	25.19

$F(z)$	Number of degrees of freedom									
	11	12	13	14	15	16	17	18	19	20
0.005	2.60	3.07	3.57	4.07	4.60	5.14	5.70	6.26	6.84	7.43
0.01	3.05	3.57	4.11	4.66	5.23	5.81	6.41	7.01	7.63	8.26
0.025	3.82	4.40	5.01	5.63	6.26	6.91	7.56	8.23	8.91	9.59
0.05	4.57	5.23	5.89	6.57	7.26	7.96	8.67	9.39	10.12	10.85
0.95	19.68	21.03	22.36	23.68	25.00	26.30	27.59	28.87	30.14	31.41
0.975	21.92	23.34	24.74	26.12	27.49	28.85	30.19	31.53	32.85	34.17
0.99	24.73	26.22	27.69	29.14	30.58	32.00	33.41	34.81	36.19	37.57
0.995	26.76	28.30	29.82	31.32	32.80	34.27	35.72	37.16	38.58	40.00

$F(z)$	Number of degrees of freedom									
	21	22	23	24	25	26	27	28	29	30
0.005	8.0	8.6	9.3	9.9	10.5	11.2	11.8	12.5	13.1	13.8
0.01	8.9	9.5	10.2	10.9	11.5	12.2	12.9	13.6	14.3	15.0
0.025	10.3	11.0	11.7	12.4	13.1	13.8	14.6	15.3	16.0	16.8
0.05	11.6	12.3	13.1	13.8	14.6	15.4	16.2	16.9	17.7	18.5
0.95	32.7	33.9	35.2	36.4	37.7	38.9	40.1	41.3	42.6	43.8
0.975	35.5	36.8	38.1	39.4	40.6	41.9	43.2	44.5	45.7	47.0
0.99	38.9	40.3	41.6	43.0	44.3	45.6	47.0	48.3	49.6	50.9
0.995	41.4	42.8	44.2	45.6	46.9	48.3	49.6	51.0	52.3	53.7

$F(z)$	Number of degrees of freedom							
	40	50	60	70	80	90	100	> 100 (Approximation)
0.005	20.7	28.0	35.5	43.3	51.2	59.2	67.3	$\frac{1}{2}(h - 2.58)^2$
0.01	22.2	29.7	37.5	45.4	53.5	61.8	70.1	$\frac{1}{2}(h - 2.33)^2$
0.025	24.4	32.4	40.5	48.8	57.2	65.6	74.2	$\frac{1}{2}(h - 1.96)^2$
0.05	26.5	34.8	43.2	51.7	60.4	69.1	77.9	$\frac{1}{2}(h - 1.64)^2$
0.95	55.8	67.5	79.1	90.5	101.9	113.1	124.3	$\frac{1}{2}(h + 1.64)^2$
0.975	59.3	71.4	83.3	95.0	106.6	118.1	129.6	$\frac{1}{2}(h + 1.96)^2$
0.99	63.7	76.2	88.4	100.4	112.3	124.1	135.8	$\frac{1}{2}(h + 2.33)^2$
0.995	66.8	79.5	92.0	104.2	116.3	128.3	140.2	$\frac{1}{2}(h + 2.58)^2$

In the last column, $h = \sqrt{2m} - 1$ where m is the number of degrees of freedom.

Table 54. F-Distribution with (m, n) Degrees of Freedom

Values of z for which the distribution function $F(z)$ (cf. (13), Sec. 18.15) has the value 0.95.
Example: For $(7, 4)$ degrees of freedom, $z = 6.09$ when $F(z) = 0.95$.

n	$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$	$m = 6$	$m = 7$	$m = 8$	$m = 9$
1	161	200	216	225	230	234	237	239	241
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
32	4.15	3.30	2.90	2.67	2.51	2.40	2.31	2.24	2.19
34	4.13	3.28	2.88	2.65	2.49	2.38	2.29	2.23	2.17
36	4.11	3.26	2.87	2.63	2.48	2.36	2.28	2.21	2.15
38	4.10	3.24	2.85	2.62	2.46	2.35	2.26	2.19	2.14
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.02
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	2.00
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97
150	3.90	3.06	2.66	2.43	2.27	2.16	2.07	2.00	1.94
200	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93
1000	3.85	3.00	2.61	2.38	2.22	2.11	2.02	1.95	1.89
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

Table 54. F-Distribution with (m, n) Degrees of Freedom (continued)Values of z for which the distribution function $F(z)$ (cf. (13), Sec. 18.15) has the value 0.95.

n	$m = 10$	$m = 15$	$m = 20$	$m = 30$	$m = 40$	$m = 50$	$m = 100$	∞
1	242	246	248	250	251	252	253	254
2	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5
3	8.79	8.70	8.66	8.62	8.59	8.58	8.55	8.53
4	5.96	5.86	5.80	5.75	5.72	5.70	5.66	5.63
5	4.74	4.62	4.56	4.50	4.46	4.44	4.41	4.37
6	4.06	3.94	3.87	3.81	3.77	3.75	3.71	3.67
7	3.64	3.51	3.44	3.38	3.34	3.32	3.27	3.23
8	3.35	3.22	3.15	3.08	3.04	3.02	2.97	2.93
9	3.14	3.01	2.94	2.86	2.83	2.80	2.76	2.71
10	2.98	2.85	2.77	2.70	2.66	2.64	2.59	2.54
11	2.85	2.72	2.65	2.57	2.53	2.51	2.46	2.40
12	2.75	2.62	2.54	2.47	2.43	2.40	2.35	2.30
13	2.67	2.53	2.46	2.38	2.34	2.31	2.26	2.21
14	2.60	2.46	2.39	2.31	2.27	2.24	2.19	2.13
15	2.54	2.40	2.33	2.25	2.20	2.18	2.12	2.07
16	2.49	2.35	2.28	2.19	2.15	2.12	2.07	2.01
17	2.45	2.31	2.23	2.15	2.10	2.08	2.02	1.96
18	2.41	2.27	2.19	2.11	2.06	2.04	1.98	1.92
19	2.38	2.23	2.16	2.07	2.03	2.00	1.94	1.88
20	2.35	2.20	2.12	2.04	1.99	1.97	1.91	1.84
22	2.30	2.15	2.07	1.98	1.94	1.91	1.85	1.78
24	2.25	2.11	2.03	1.94	1.89	1.86	1.80	1.73
26	2.22	2.07	1.99	1.90	1.85	1.82	1.76	1.69
28	2.19	2.04	1.96	1.87	1.82	1.79	1.73	1.65
30	2.16	2.01	1.93	1.84	1.79	1.76	1.70	1.62
32	2.14	1.99	1.91	1.82	1.77	1.74	1.67	1.59
34	2.12	1.97	1.89	1.80	1.75	1.71	1.65	1.57
36	2.11	1.95	1.87	1.78	1.73	1.69	1.62	1.55
38	2.09	1.94	1.85	1.76	1.71	1.68	1.61	1.53
40	2.08	1.92	1.84	1.74	1.69	1.66	1.59	1.51
50	2.03	1.87	1.78	1.69	1.63	1.60	1.52	1.44
60	1.99	1.84	1.75	1.65	1.59	1.56	1.48	1.39
70	1.97	1.81	1.72	1.62	1.57	1.53	1.45	1.35
80	1.95	1.79	1.70	1.60	1.54	1.51	1.43	1.32
90	1.94	1.78	1.69	1.59	1.53	1.49	1.41	1.30
100	1.93	1.77	1.68	1.57	1.52	1.48	1.39	1.28
150	1.89	1.73	1.64	1.53	1.48	1.44	1.34	1.22
200	1.88	1.72	1.62	1.52	1.46	1.41	1.32	1.19
1000	1.84	1.68	1.58	1.47	1.41	1.36	1.26	1.08
∞	1.83	1.67	1.57	1.46	1.39	1.35	1.24	1.00

Table 54. F-Distribution with (m, n) Degrees of Freedom (continued)Values of z for which the distribution function $F(z)$ (cf. (13), Sec. 18.15) has the value 0.99

n	$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$	$m = 6$	$m = 7$	$m = 8$	$m = 9$
1	4052	4999	5403	5625	5764	5859	5928	5982	6022
2	98.5	99.0	99.2	99.3	99.3	99.3	99.4	99.4	99.4
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
14	8.86	6.51	5.56	5.04	4.70	4.46	4.28	4.14	4.03
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
32	7.50	5.34	4.46	3.97	3.65	3.43	3.26	3.13	3.02
34	7.44	5.29	4.42	3.93	3.61	3.39	3.22	3.09	2.98
36	7.40	5.25	4.38	3.89	3.57	3.35	3.18	3.05	2.95
38	7.35	5.21	4.34	3.86	3.54	3.32	3.15	3.02	2.92
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
50	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.79
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
70	7.01	4.92	4.08	3.60	3.29	3.07	2.91	2.78	2.67
80	6.96	4.88	4.04	3.56	3.26	3.04	2.87	2.74	2.64
90	6.93	4.85	4.01	3.54	3.23	3.01	2.84	2.72	2.61
100	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59
150	6.81	4.75	3.92	3.45	3.14	2.92	2.76	2.63	2.53
200	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50
1000	6.66	4.63	3.80	3.34	3.04	2.82	2.66	2.53	2.43
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41

Table 54. F-Distribution with (m, n) Degrees of Freedom (continued)Values of z for which the distribution function $F(z)$ (cf. (13), Sec. 18.15) has the value 0.99.

n	$m = 10$	$m = 15$	$m = 20$	$m = 30$	$m = 40$	$m = 50$	$m = 100$	∞
1	6056	6157	6209	6261	6287	6300	6330	6366
2	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5
3	27.2	26.9	26.7	26.5	26.4	26.4	26.2	26.1
4	14.5	14.2	14.0	13.8	13.7	13.7	13.6	13.5
5	10.1	9.72	9.55	9.38	9.29	9.24	9.13	9.02
6	7.87	7.56	7.40	7.23	7.14	7.09	6.99	6.88
7	6.62	6.31	6.16	5.99	5.91	5.86	5.75	5.65
8	5.81	5.52	5.36	5.20	5.12	5.07	4.96	4.86
9	5.26	4.96	4.81	4.65	4.57	4.52	4.42	4.31
10	4.85	4.56	4.41	4.25	4.17	4.12	4.01	3.91
11	4.54	4.25	4.10	3.94	3.86	3.81	3.71	3.60
12	4.30	4.01	3.86	3.70	3.62	3.57	3.47	3.36
13	4.10	3.82	3.66	3.51	3.43	3.38	3.27	3.17
14	3.94	3.66	3.51	3.35	3.27	3.22	3.11	3.00
15	3.80	3.52	3.37	3.21	3.13	3.08	2.98	2.87
16	3.69	3.41	3.26	3.10	3.02	2.97	2.86	2.75
17	3.59	3.31	3.16	3.00	2.92	2.87	2.76	2.65
18	3.51	3.23	3.08	2.92	2.84	2.78	2.68	2.57
19	3.43	3.15	3.00	2.84	2.76	2.71	2.60	2.49
20	3.37	3.09	2.94	2.78	2.69	2.64	2.54	2.42
22	3.26	2.98	2.83	2.67	2.58	2.53	2.42	2.31
24	3.17	2.89	2.74	2.58	2.49	2.44	2.33	2.21
26	3.09	2.82	2.66	2.50	2.42	2.36	2.25	2.13
28	3.03	2.75	2.60	2.44	2.35	2.30	2.19	2.06
30	2.98	2.70	2.55	2.39	2.30	2.25	2.13	2.01
32	2.93	2.66	2.50	2.34	2.25	2.20	2.08	1.96
34	2.89	2.62	2.46	2.30	2.21	2.16	2.04	1.91
36	2.86	2.58	2.43	2.26	2.17	2.12	2.00	1.87
38	2.83	2.55	2.40	2.23	2.14	2.09	1.97	1.84
40	2.80	2.52	2.37	2.20	2.11	2.06	1.94	1.80
50	2.70	2.42	2.27	2.10	2.01	1.95	1.82	1.68
60	2.63	2.35	2.20	2.03	1.94	1.88	1.75	1.60
70	2.59	2.31	2.15	1.98	1.89	1.83	1.70	1.54
80	2.55	2.27	2.12	1.94	1.85	1.79	1.66	1.49
90	2.52	2.24	2.09	1.92	1.82	1.76	1.62	1.46
100	2.50	2.22	2.07	1.89	1.80	1.73	1.60	1.43
150	2.44	2.16	2.00	1.83	1.73	1.66	1.52	1.33
200	2.41	2.13	1.97	1.79	1.69	1.63	1.48	1.28
1000	2.34	2.06	1.90	1.72	1.61	1.54	1.38	1.11
∞	2.32	2.04	1.88	1.70	1.59	1.52	1.36	1.00

Table 55. Distribution Function $F(x) = P(I \leq x)$ of the Random Variable I in Section 18.19

When $n = 3$, $F(2) = 1 - 0.167 = 0.833$.

When $n = 4$, $F(3) = 1 - 0.375 = 0.625$, $F(4) = 1 - 0.167 = 0.833$, etc.

x	$n = 3$	x	$n = 4$	x	$n = 5$	x	$n = 6$	x	$n = 7$	x	$n = 8$	x	$n = 9$	x	$n = 10$	x	$n = 11$
0	0.167	0	0.042	0	0.008	0	0.001	1	0.001	2	0.001	4	0.001	6	0.001	8	0.001
1	500	1	167	1	42	1	8	2	0.005	3	0.003	5	0.003	7	0.002	9	0.002
2	375	2	117	2	28	3	0.068	4	0.035	5	0.016	6	0.006	8	0.005	10	0.003
3	242	3	42	4	136	5	0.68	6	0.31	7	0.12	9	0.008	11	0.005	13	0.003
4	408	5	235	6	119	7	54	8	0.89	10	0.60	12	0.36	14	0.20	16	0.08
5	360	6	191	7	281	9	138	10	199	11	0.90	13	0.54	15	0.30	17	0.008
6	500	8	386	9	500	10	500	11	274	12	130	14	0.78	16	0.43	18	0.002
7	500	10	500	11	274	12	360	13	179	15	108	17	100	19	109	21	0.001
8	500	12	360	13	452	14	238	15	306	17	190	18	242	20	141	22	223
9	500	13	452	14	452	15	306	16	381	17	460	19	300	21	271	23	324
10	500	14	452	15	306	16	381	17	460	19	300	20	364	21	431	23	381
11	500	15	306	16	381	17	460	18	500	20	364	21	431	23	324	25	440
12	500	16	381	17	460	18	500	19	500	21	431	22	500	24	381	26	440
13	500	17	460	18	500	19	500	20	500	22	500	24	500	26	500	28	500
14	500	18	500	19	500	20	500	21	500	23	500	25	500	27	500	29	500
15	500	19	500	20	500	21	500	22	500	24	500	26	500	28	500	30	500
16	500	20	500	21	500	22	500	23	500	25	500	27	500	29	500	31	500
17	500	21	500	22	500	23	500	24	500	26	500	28	500	30	500	32	500
18	500	22	500	23	500	24	500	25	500	27	500	29	500	31	500	33	500
19	500	23	500	24	500	25	500	26	500	28	500	30	500	32	500	34	500
20	500	24	500	25	500	26	500	27	500	29	500	31	500	33	500	35	500
21	500	25	500	26	500	27	500	28	500	30	500	32	500	34	500	36	500
22	500	26	500	27	500	28	500	29	500	31	500	33	500	35	500	37	500
23	500	27	500	28	500	29	500	30	500	32	500	34	500	36	500	38	500
24	500	28	500	29	500	30	500	31	500	33	500	35	500	37	500	39	500
25	500	29	500	30	500	31	500	32	500	34	500	36	500	38	500	40	500
26	500	30	500	31	500	32	500	33	500	35	500	37	500	39	500	41	500
27	500	31	500	32	500	33	500	34	500	36	500	38	500	40	500	42	500
28	500	32	500	33	500	34	500	35	500	37	500	39	500	41	500	43	500
29	500	33	500	34	500	35	500	36	500	38	500	40	500	42	500	44	500
30	500	34	500	35	500	36	500	37	500	39	500	41	500	43	500	45	500
31	500	35	500	36	500	37	500	38	500	40	500	42	500	44	500	46	500
32	500	36	500	37	500	38	500	39	500	41	500	43	500	45	500	47	500
33	500	37	500	38	500	39	500	40	500	42	500	44	500	46	500	48	500
34	500	38	500	39	500	40	500	41	500	43	500	45	500	47	500	49	500
35	500	39	500	40	500	41	500	42	500	44	500	46	500	48	500	50	500
36	500	40	500	41	500	42	500	43	500	45	500	47	500	49	500	51	500
37	500	41	500	42	500	43	500	44	500	46	500	48	500	50	500	52	500
38	500	42	500	43	500	44	500	45	500	47	500	49	500	51	500	53	500
39	500	43	500	44	500	45	500	46	500	48	500	50	500	52	500	54	500
40	500	44	500	45	500	46	500	47	500	49	500	51	500	53	500	55	500
41	500	45	500	46	500	47	500	48	500	50	500	52	500	54	500	56	500
42	500	46	500	47	500	48	500	49	500	51	500	53	500	55	500	57	500
43	500	47	500	48	500	49	500	50	500	52	500	54	500	56	500	58	500
44	500	48	500	49	500	50	500	51	500	53	500	55	500	57	500	59	500
45	500	49	500	50	500	51	500	52	500	54	500	56	500	58	500	60	500
46	500	50	500	51	500	52	500	53	500	55	500	57	500	59	500	61	500
47	500	51	500	52	500	53	500	54	500	56	500	58	500	60	500	62	500
48	500	52	500	53	500	54	500	55	500	57	500	59	500	61	500	63	500
49	500	53	500	54	500	55	500	56	500	58	500	60	500	62	500	64	500
50	500	54	500	55	500	56	500	57	500	59	500	61	500	63	500	65	500
51	500	55	500	56	500	57	500	58	500	60	500	62	500	64	500	66	500
52	500	56	500	57	500	58	500	59	500	61	500	63	500	65	500	67	500
53	500	57	500	58	500	59	500	60	500	62	500	64	500	66	500	68	500
54	500	58	500	59	500	60	500	61	500	63	500	65	500	67	500	69	500
55	500	59	500	60	500	61	500	62	500	64	500	66	500	68	500	70	500
56	500	60	500	61	500	62	500	63	500	65	500	67	500	69	500	71	500
57	500	61	500	62	500	63	500	64	500	66	500	68	500	70	500	72	500
58	500	62	500	63	500	64	500	65	500	67	500	69	500	71	500	73	500
59	500	63	500	64	500	65	500	66	500	68	500	70	500	72	500	74	500
60	500	64	500	65	500	66	500	67	500	69	500	71	500	73	500	75	500
61	500	65	500	66	500	67	500	68	500	70	500	72	500	74	500	76	500
62	500	66	500	67	500	68	500	69	500	71	500	73	500	75	500	77	500
63	500	67	500	68	500	69	500	70	500	72	500	74	500	76	500	78	500
64	500	68	500	69	500	70	500	71	500	73	500	75	500	77	500	79	500
65	500	69	500	70	500	71	500	72	500	74	500	76	500	78	500	80	500
66	500	70	500	71	500	72	500	73	500	75	500	77	500	79	500	81	500
67	500	71	500	72	500	73	500	74	500	76	500	78	500	80	500	82	500
68	500	72	500	73	500	74	500	75	500	77	500	79	500	81	500	83	500
69	500	73	500	74	500	75	500	76	500	78	500	80	500	82	500	84	500
70	500	74	500	75	500	76	500	77	500	79	500	81	500	83	500	85	500
71	500	75	500	76	500	77	500	78	500	80	500	82	500	84	500	86	500
72	500	76	500	77	500	78	500	79	500	81	500	83	500	85	500	87	500
73	500	77	500	78	500	79	500	80	500	82	500	84	500	86	500	88	500
74	500	78	500	79	500	80	500	81	500	83	500	85	500	87	500	89	500
75	500	79	500	80	500	81	500	82	500	84	500	86	500	88	500	90	500
76	500	80	500	81	500	82	500	83	500	85	500	87	500	89	500	91	500
77	500	81	500	82	500	83	500	84	500	86	500	88	500	90	500	92	500
78	500	82	500	83	500	84	500	85	500	87	500	89	500	91	500	93	500
79	500	83	500	84	500	85	500	86	500	88	500	90	500	92	500	94	500
80	500	84	500	85	500	86	500	87	500	89	500	91	500	93	500	95	500
81	500	85	500	86	500	87	500	88	500	90	500	92	500	94	500	96	500
82	50																

Appendix I

References

A General References

- [A1] Courant, R., *Differential and Integral Calculus*. 2 vols. New York: Interscience, 1964, 1956.
- [A2] Courant, R., and D. Hilbert, *Methods of Mathematical Physics*. New York: Interscience, 1955.
- [A3] C.R.C. *Standard Mathematical Tables*. 11th ed. Cleveland: Chemical Rubber Pub. Co., 1957.
- [A4] Dwight, H. B., *Mathematical Tables*. New York: McGraw-Hill, 1941.
- [A5] Erdélyi, A., W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions*. 3 vols. New York: McGraw-Hill, 1953, 1955.
- [A6] Fletcher, A., J. C. P. Miller, L. Rosenhead, and L. J. Comrie, *An Index of Mathematical Tables*. Oxford: Blackwell, 1962.
- [A7] Jahnke, E., F. Emde, and F. Lösch, *Tables of Higher Functions*. 7th ed. Stuttgart: Teubner, 1966.
- [A8] Kaplan, W., *Advanced Calculus*. 5th printing. Reading, Mass.: Addison-Wesley, 1959.
- [A9] Knopp, K., *Infinite Sequences and Series*. New York: Dover, 1956.
- [A10] Magnus, W., and F. Oberhettinger, *Formulas and Theorems for the Special Functions of Mathematical Physics*, New York: Chelsea, 1949.
- [A11] Milne, W. E., *Numerical Calculus*. Princeton, N.J.: Princeton University Press, 1949.
- [A12] Taylor, A. E., *Advanced Calculus*. New York: Ginn, 1955.
- [A13] Thomas, G. B., *Calculus and Analytic Geometry*. 2nd ed., 5th printing. Reading, Mass.: Addison-Wesley, 1953.
- [A14] Whittaker, E. T., and G. N. Watson, *A Course of Modern Analysis*. New York: Macmillan, 1947.
- [A15] Willers, F. A., *Practical Analysis*. New York: Dover, 1948.

B Ordinary Differential Equations (Chaps. 1-4)

- [A2], [A5], [A10], [A15].
- [B1] Carslaw, H. S., and J. C. Jaeger, *Operational Methods in Applied Mathematics*. 2nd ed. London: Oxford University Press, 1948.
- [B2] Churchill, R. V., *Operational Mathematics*. New York: McGraw-Hill, 1958.
- [B3] Coddington, E. A., and N. Levinson, *Theory of Ordinary Differential Equations*. New York: McGraw-Hill, 1955.
- [B4] Forsyth, A. R., *A Treatise on Differential Equations*. 6th ed. London: Macmillan, 1951.
- [B5] Ince, E. L., *Ordinary Differential Equations*. New York: Dover, 1944.

- [B6] Kamke, E., *Differentialgleichungen, Lösungsmethoden und Lösungen. I. Gewöhnliche Differentialgleichungen*. 6th ed. Leipzig: Akademische Verlagsgesellschaft, 1959. (This extremely useful book contains a systematic list of more than 1500 differential equations and their solutions.)
- [B7] Levy, H., and E. A. Baggott, *Numerical Solutions of Differential Equations*. New York: Dover, 1950.
- [B8] Morris, M., and O. E. Brown, *Differential Equations*. 3rd ed. New York: Prentice-Hall, 1952.
- [B9] Watson, G. N., *A Treatise on the Theory of Bessel Functions*. 2nd ed. New York: Macmillan, 1944.
- [B10] Widder, D. V., *The Laplace Transform*. Princeton, N.J.: Princeton University Press, 1941.

Tables

[A3].

- [B11] Erdélyi, A., W. Magnus, F. Oberhettinger, and F. Tricomi, *Tables of Integral Transforms*. 2 vols. New York: McGraw-Hill, 1954.

C Vector Analysis and Integral Theorems (Chaps. 5 and 6)

[A1], [A8], [A12].

- [C1] Brand, L., *Vector Analysis*. 3rd printing. New York: Wiley, 1961.

[C2] Brand, L., *Vectorial Mechanics*. New York: Wiley, 1930.[C3] Kellogg, O. D., *Foundations of Potential Theory*. New York: Dover, 1953.[C4] Kreyszig, E., *Introduction to Differential Geometry and Riemannian Geometry*. Toronto: Toronto University Press, 1967.[C5] Lamb, H., *Hydrodynamics*. 6th ed. New York: Dover, 1945.[C6] Struik, D. J., *Lectures on Classical Differential Geometry*. Reading, Mass.: Addison-Wesley, 1950.[C7] Weatherburn, C. E., *Elementary Vector Analysis with Applications to Geometry and Physics*. London: Bell, 1921.[C8] Weatherburn, C. E., *Advanced Vector Analysis with Applications to Mathematical Physics*. London: Bell, 1928.**D Matrices (Chap. 7)**[D1] Bodewig, E., *Matrix Calculus*. 2nd ed. Amsterdam: North-Holland Pub. 1959.[D2] Frazer, R. A., W. J. Duncan, and A. R. Collar, *Elementary Matrices*. Cambridge: University Press, 1938.[D3] Gantmacher, F. R., *The Theory of Matrices*. 2 vols. New York: Chelsea, 1959.[D4] McDuffee, C. C., *The Theory of Matrices*. New York: Chelsea, 1946.[D5] Perlis, S., *Theory of Matrices*. Reading, Mass.: Addison-Wesley, 1952.[D6] Schreier, O., and E. Spencer, *Introduction to Modern Algebra and Matrix Theory*. New York: Chelsea, 1951.[D7] Schwerdtfeger, H., *Introduction to Linear Algebra and the Theory of Matrices*. Groningen: Noordhoff, 1950.**E Fourier Series and Integrals (Chap. 8)**

[A9], [A14], [A15].

[E1] Carslaw, H. S., *Introduction to the Theory of Fourier's Series and Integrals*. 3rd ed. London: Macmillan, 1930.[E2] Churchill, R. V., *Fourier Series and Boundary Value Problems*. New York: McGraw-Hill, 1941.[E3] Rogosinski, W., *Fourier Series*. New York: Chelsea, 1950.[E4] Sneddon, I. N., *Fourier Transforms*. New York: McGraw-Hill, 1951.[E5] Szegő, G., *Orthogonal Polynomials*. Revised ed. New York: American Mathematical Society, 1959.

- [E6] Titchmarsh, E. C., *Introduction to the Theory of Fourier Integrals*. Oxford: Clarendon, 1948.
- [E7] Zygmund, A., *Trigonometric Series*. 2 vols. Cambridge: University Press, 1959.
- Tables of Fourier Transforms*
- [B11], [E4].
- [E8] Campbell, G. A., and R. M. Foster, *Fourier Integrals for Practical Applications*. Bell Telephone System Tech. Pub., 1942.

F Partial Differential Equations (Chap. 9)

- [A2], [A14], [C3], [C5], [E2].
- [F1] Bateman, H., *Partial Differential Equations of Mathematical Physics*. New York: Dover, 1944.
- [F2] Bergman, S., and M. Schiffer, *Kernel Functions and Elliptic Differential Equations in Mathematical Physics*. New York: Academic Press, 1953.
- [F3] Duff, G. F. D., *Partial Differential Equations*. Toronto: Toronto University Press, 1956.
- [F4] Lord Rayleigh, *The Theory of Sound*. 2 vols. New York: Dover, 1945.
- [F5] Sagan, H., *Boundary and Eigenvalue Problems in Mathematical Physics*. New York: Wiley, 1961.
- [F6] Sneddon, I. N., *Elements of Partial Differential Equations*. New York: McGraw-Hill, 1957.
- [F7] Sommerfeld, A., *Partial Differential Equations in Physics*. New York: Academic Press, 1949.
- [F8] Webster, A. G., *Partial Differential Equations of Mathematical Physics*. 2nd ed. New York: Hafner, 1947.

G Complex Analysis (Chaps. 10-16)

- [A9], [A14], [C5].
- [G1] Ahlfors, L. V., *Complex Analysis*. New York: McGraw-Hill, 1953.
- [G2] Bierbach, L., *Conformal Mapping*. New York: Chelsea, 1953.
- [G3] Churchill, R. V., *Complex Variables and Applications*. 2nd ed. New York: McGraw-Hill, 1960.
- [G4] Hille, E., *Analytic Function Theory*. 2 vols. Boston: Ginn, 1959, 1962.
- [G5] Knopp, K., *Theory of Functions*. 2 vols. New York: Dover, 1945.
- [G6] Nehari, Z., *Conformal Mapping*. New York: McGraw-Hill, 1952.
- [G7] Rothe, R., F. Ollendorf, and K. Pohlhausen, *Theory of Functions as Applied to Engineering Problems*. Cambridge, Mass.: Technology Press, 1933.
- [G8] Springer, G., *Introduction to Riemann Surfaces*. Reading, Mass.: Addison-Wesley, 1957.
- [G9] Titchmarsh, E. C., *The Theory of Functions*. 2nd ed. London: Oxford University Press, 1939.

H Special Functions. Asymptotic Expansions (Chap. 17)

- [A3]-[A7], [A9], [A10], [A14], [G5].
- [H1] Borel, E., *Leçons sur les séries divergentes*. 2nd ed. Paris: Gauthier-Villars, 1928.
- [H2] Erdélyi, A., *Asymptotic Expansions*. New York: Dover, 1956.
- [H3] Rainville, E. D., *Special Functions*. New York: Macmillan, 1960.

I Probability and Statistics (Chap. 18)

- [I1] Cochran, W. G., *Sampling Techniques*. 2nd ed. New York: Wiley, 1963.
- [I2] Cox, D. R., *Planning of Experiments*. New York: Wiley, 1958.
- [I3] Cramér, H., *Mathematical Methods of Statistics*. 8th printing. Princeton: Princeton University Press, 1963..

- [I4] Feller, W., *An Introduction to Probability and Its Applications*. Vol. I. 2nd ed., 6th printing. New York: Wiley, 1961.
- [I5] Fisher, R. A., *Statistical Methods for Research Workers*. 13th ed. New York: Hafner, 1958.
- [I6] Grant, E. L., *Statistical Quality Control*. New York: McGraw-Hill, 1952.
- [I7] Huff, D., *How to Lie with Statistics*. New York: Norton, 1954.
- [I8] Rice, W. B., *Control Charts in Factory Management*. New York: Wiley, 1947.
- [I9] Wilks, S. S., *Mathematical Statistics*. New York: Wiley, 1962.

J Tables

- [A3], [A4], [A7].
- [J1] British Association for the Advancement of Science, *Mathematical Tables*. Vol. VII. *The Probability Integral*. Cambridge: University Press, 1939.
- [J2] Davis, H. T., *Tables of Higher Mathematical Functions*. 2 vols. Bloomington, Ind.: Principia Press, 1933.
- [J3] Dodge, H. F., and H. G. Romig, *Sampling Inspection Tables*. 2nd ed. New York: Wiley, 1959.
- [J4] Fisher, R. A., and F. Yates, *Statistical Tables for Biological, Agricultural, and Medical Research*. 5th ed. London: Oliver and Boyd, 1957.
- [J5] Hald, A., *Statistical Tables and Formulas*. 4th printing. New York: Wiley, 1962.
- [J6] Harvard Computation Laboratory, *Tables of the Cumulative Binomial Probability Distribution*. Cambridge, Mass.: Harvard University Press, 1955.
- [J7] National Bureau of Standards, *Tables of Normal Probability Functions*. Washington, D.C.: US Gov. Printing Office, 1953.
- [J8] National Bureau of Standards, *Tables of the Error Function and Its Derivative*. Washington, D.C.: US Gov. Printing Office, 1954.
- [J9] Pearson, E. S., and H. O. Hartley, *Biometrika Tables for Statisticians*. Vol. I. 2nd ed. Cambridge: University Press, 1958.
- [J10] Rand Corporation, *A Million Random Digits with 100,000 Normal Deviates*. Glencoe: Free Press, 1955.

Appendix 2

Answers to Odd-Numbered Problems

Remark. In the case of more complicated problems, we shall present not only the answer but also the solution. Odd-numbered problems not occurring in this Appendix are those in which obvious graphical work, verifications of statements and formulas in the text, or simple proofs are required, or the answer is already given in the enunciation of the problem.

SECTION 1.1

1. First order

$$13. y' = 1 + y^2$$

$$19. y = 3e^{-x}$$

3. Second order

$$15. y = (v^3 - 8)/3$$

$$21. y = -2e^{-x^2}$$

$$11. y' = 3y/x$$

$$17. y = \frac{1}{2} \sin 2x + 1$$

$$23. y = 0.01e^{kt}$$

SECTION 1.2

19. The last integral has the value 0.75, approximately. It cannot be evaluated by elementary methods and will be considered later (Sec. 17.2).

SECTION 1.3

3. Suppose you forgot \tilde{c} in (4), wrote $\ln y = -x^2$, transformed this to $y = e^{-x^2}$, and then afterwards added a constant c , obtaining $y = e^{-x^2} + c$, which is not a solution of $y' = -2xy$.

$$5. y = cx^a$$

$$11. y^3 + x^3 = c$$

$$17. y = c\sqrt{\sin 2x}$$

$$23. y = y_0 \cosh x$$

$$29. y = (\ln x)^2$$

$$33. y^a = x/2 - (\sin 2wx)/4w + 1$$

37. 25 min

39. $k = (h^{3/4} - 2Rh^{1/4})h'$ where $k = 0.6r_0^{2/3}\sqrt{2g}$, $kt + c = 2h^{3/4}/5 - 4Rh^{3/4}/3$; $h(0) = R$ gives $c = -14R^{3/4}/15$; to $h = 0$ there corresponds $t = 1400/1.8\sqrt{g} \approx 25$ [sec].

41. $y' = ky$, $y = y_0 e^{kt}$. At $t = 1590$, $y = y_0/2$. Consequently,

$$1590k = \ln \frac{1}{2}, \quad k = -0.000436, \quad y(1) = y_0 e^{-0.000436}.$$

For small $|x|$, $e^x \approx 1 + x$. Thus, $y(1) \approx y_0(1 - 0.000436)$, $y_0 - y(1) \approx 0.0004$, that is, 0.04% will disappear.

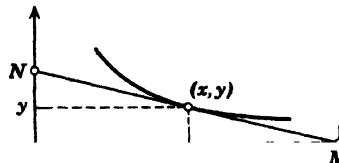
43. $y_0 = 4$ g

47. $\Delta A = -kA \Delta x$ (A = amount of incident light, ΔA = absorbed light, Δx = thickness, $-k$ = constant of proportionality). Let $\Delta x \rightarrow 0$. Then $A' = -kA$, so $A(x) = A_0 e^{-kx}$ is the amount of light in a thick layer at depth x from the surface of incidence.

49. $y' = -0.05y$, $y(0) = 10$, $y = 10e^{-0.05t}$, $y(120) = 0.025$

51. $x^2 + y^2 = 25$

53. Let $y(x)$ be such a curve. Then $OM = 2x$, $ON = 2y$, and the slope at (x, y) is $y' = -ON/OM = -y/x$. Thus $xy = c$.



Problem 53

55. Since y' is the slope of the tangent, the slope a of the normal $y = ax$ passing through the origin is $a = -1/y'$. Thus $y = -x/y'$, $x^2 + y^2 = c$ (circles).

SECTION 1.4

1. $y = x(\ln x + c)$

3. $y = cx^{-1} - x$

5. $y = x + x/(c - \ln x)$

7. $y = x \arcsin(x + c)$

9. $y = x + x(c - x^2)^{-\frac{1}{2}}$

11. $y = x + c(x + 1)^2$

13. Put $y/x = u$, $\ln(x^2 + y^2) + 2 \arctan \frac{y}{x} = c$

15. $\cos(y + x) = ce^{-x}$

17. $r = l \exp(-\theta - \frac{1}{4}\pi)/\sqrt{2}$, etc., where l is the length of the table and r and θ are polar coordinates with center at the center of the table. Graph this logarithmic spiral.

SECTION 1.5

1. $du = [2x + 2y^2 - y \cos(xy)] dx + [4xy - x \cos(xy)] dy$

3. $du = [e^x \cos y + 2(x - y)] dx - [e^x \sin y + 2(x - y)] dy$

5. $(3x^2y + \cos x) dx + x^3 dy = 0$

7. $e^{-2xy}((1 - 2xy) dx - 2x^2 dy) = 0$

9. $x^2 + 4y^2 = c$

11. Not exact, $y = 1/(c + 2 \ln x)$

13. Not exact, $x^4 \sin y = c$

15. $(x^3 + \ln x) y = c$

17. $e^{xy} + y^2 = c$

19. $(x - 3)^2 + 9(y - 1)^2 = 9$

21. $y = \tan x$

23. Circle $(x - 1)^2 + (y + 1)^2 = 1$

25. Hyperbola $(x + 2)^2 - (y + 1)^2 = 1$

31. $b = a$, $x^2 + axy + y^2 = c$

33. $f(x) dg/dy = 0$, hence $g = \text{const}$

SECTION 1.6

1. $F = xy$, $x^2y^3 = c$

3. $F = 1/x$, $y = c/\ln x$

5. $F = 1/y^3$, $y = c\sqrt{x}$

7. $F = 1/x^2$, $\sin y = cx$

9. $F = x$, $x^2y = -4$

11. $F = e^x$, $e^x \cos y = -1$

13. $\frac{\partial FP}{\partial y} = \frac{\partial FQ}{\partial x} \cdot \frac{k}{y} P + \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$

15. $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} + Q$

17. $\frac{\partial P}{\partial y} + \frac{k}{y} P = \frac{\partial Q}{\partial x} + \frac{m}{x} Q$

19. $\frac{\partial FP}{\partial y} = \frac{\partial FQ}{\partial x} \cdot \frac{\partial F}{\partial y} P + F \frac{\partial P}{\partial y} - \frac{\partial F}{\partial x} Q - F \frac{\partial Q}{\partial x} = 0$. Choosing $F = F(x)$, using the assumption, and dividing by Q , we obtain $Ff - F' = 0$. $F = e^{\int f dx}$.

SECTION 1.7

✓3. $y = ce^x + x$

✓5. $y = ce^{4x} + x^2$

7. $y = ce^{x^2} + x$

9. $y = ce^{3x} - \frac{2}{13} \cos 2x - \frac{3}{13} \sin 2x$

11. $y = cx^{-1} + 0.25x^3 + 0.5x$

15. $y = ce^{-3x} + 0.2e^{2x} + 2$

19. $y = -2x^2 - x - 0.25$

23. $y_1' + fy_1 = 0$, $y_2' + fy_2 = 0$, hence $(y_1 + y_2)' + f(y_1 + y_2) = 0$.

25. $y_1' + fy_1 = r$, $y_2' + fy_2 = r$, hence $(y_1 - y_2)' + f(y_1 - y_2) = r - r = 0$.

27. $y_1' + fy_1 = r$, $y_2' + fy_2 = r$, hence $(y_1 + y_2)' + f(y_1 + y_2) = 2r$.

13. $y = cx^{-2} + x$

17. $y = ce^{3x} + xe^{3x}$

21. $y = 0.5(e^{-bx} + e^{bx})$

SECTION 1.8

1. $v = e^{-x}$, $u = 2e^x + c$, $y = ce^{-x} + 2$

5. $y = \sqrt{cx^2 + 1}$

9. $x = cy^2 + 0.5y^4$

13. $\cosh y = z$, $z = (c + x)e^x$

3. $y = (3x + c)e^x$

7. $y^3 = u$, $u = ce^{-x^2/2} + 1$

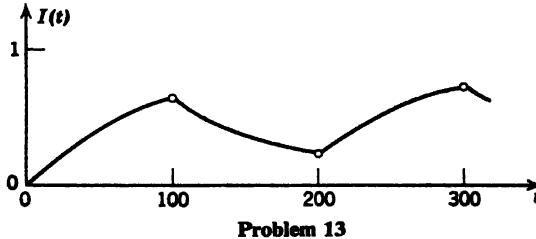
11. $\sin y = z$, $z = ce^{-x} + 5$

SECTION 1.9

7. $\tau = 1/\ln 100 \approx 0.217$ [sec], $R = L/\tau \approx 46$ [ohms]

9. $t = (\ln 1000)/250000 \approx 3 \cdot 10^{-5}$ [sec] 11. $I(0) = -E_0 \omega L / (R^2 + \omega^2 L^2)$

13. From (5*), $I = I_1 = 1 - e^{-0.01t}$ ($0 \leq t \leq 100$), $I = I_2 = c_2 e^{-0.01(t-100)}$ ($100 \leq t \leq 200$); $I_1(100) = I_2(100)$ yields $c_2 = 1 - e^{-1} \approx 0.63$. $I = I_3 = 1 + c_3 e^{-0.01(t-200)}$ ($200 \leq t \leq 300$); $I_2(200) = I_3(200)$ yields $c_3 = (1 - e^{-1})e^{-1} - 1 \approx -0.77$, etc.



Problem 13

15. $I(0) = \omega E_0 C / (1 + (\omega RC)^2)$

17. $I = 0.035(\cos 314t + 314 \sin 314t - e^{-t}) \approx 11 \sin 314t$

19. $20Q' + 100Q = 10$ (cf. Prob. 18), $Q(0) = 0$, $Q = 0.1(1 - e^{-5t})$, $V = Q/C = 10(1 - e^{-5t})$, $I = dQ/dt = 0.5e^{-5t}$, cf. (3*) and (3').

21. $I = 2 - 2(1 - 0.1t)^5$ when $0 \leq t \leq 10$, $I = 2$ when $t > 10$.

23. $I = (t - 1) \ln(1 - t)$

25. $Q = 1 - (1 - 0.1t)^{10}$ when $0 \leq t \leq 10$, $Q = 1$ when $t > 10$, $V = Q/C = 10Q$.

SECTION 1.10

1 Parallel straight lines

3. Circles of radius 2 with centers on the y -axis

5. Hyperbolas with common asymptotes 7. Congruent parabolas
 9. Congruent cubical parabolas 11. $(x - c)^2 + (y - c)^2 = 2c^2$
 13. $y + xy' = 0$ 15. $y' = 3y$ 17. $xy' = y \ln y$ 19. $y' = 3y/x$
 23. $y' = y/x$, $y' = -x/y$, $y^2 + x^2 = r^2$ 25. $y^2 + x^2/2 = c^2$
 27. $y^2 - x^2 = c^2$ 29. $y = c^* - x/2 - (\sin 2x)/4$
 31. $\frac{\partial g}{\partial x} = 0$, that is $\frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial y} dy = 0$. Thus $\frac{dy}{dx} = -\frac{\partial g/\partial x}{\partial g/\partial y}$. If we replace the right side by its negative reciprocal, the result follows.
 33. $y^2 - x^2 = c^*$ 35. $y = c^* e^{-2x}$ 37. Use (14) in Sec. 0.1.

SECTION 1.11

1. $y_0 = 1$, $y_1 = 1 + \int_0^x dt = 1 + x$, $y_2 = 1 + \int_0^x (1+t) dt = 1 + x + \frac{x^2}{2!}$, etc.
 3. $y_1(1) = 2$, $y_2(1) = 2.5$, $y_3(1) = 2.667$
 5. $y_0 = 1$, $y_n = 2\left(1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}\right) + \frac{x^{n+1}}{(n+1)!} - 1 - x$
 7. $y_0 = 0$, $y_n = \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^{n+1}}{(n+1)!}$
 9. $y_0 = 2$, $y_n = 2\left[1 + \frac{x^2}{2} + \frac{1}{2!}\left(\frac{x^2}{2}\right)^2 + \cdots + \frac{1}{n!}\left(\frac{x^2}{2}\right)^n\right]$
 11. $y_0 = 0$, $y_1 = x^2 - \frac{x^4}{4}$, $y_2 = x^2 - \frac{x^6}{4 \cdot 6}$, $y_3 = x^2 - \frac{x^8}{4 \cdot 6 \cdot 8}$, etc.
 13. $y_0 = 0$, $y_1 = x$, $y_2 = x + \frac{x^3}{3}$, $y_3 = x + \frac{x^3}{3} + \frac{x^5}{3 \cdot 5}$, ...

SECTION 1.13

7. 0.1, 0.202, 0.31008, 0.42869 etc. Exact solution: $y = e^{x^2} \int_0^x e^{-t^2} dt$

11. The Runge-Kutta method is a fourth-order method.

13. $y = \tan x$

SECTION 2.1

9. $z = ce^{-4x}$, $y = c_1 e^{-4x} + c_2$ 11. $y = c_1 x^{-1} + c_2$
 13. $y = -\ln \cos(x + c_1) + c_2$
 15. $yz(dx/dy) = 2z^2$, $z = cy^2 = y'$, $y = 1/(c_1 x + c_2)$
 17. $y = \ln(c_1 x + c_2)$
 21. $y = -y^2$, $y(t) = 1 + \ln(3t + 1)$, $y(5) = 1 + \ln 16 \approx 5.77$ [ft],
 $\dot{y}(5) = \frac{3}{16} \approx 0.19$ [ft/sec]
 23. $y(t) = (2t)^{3/2}/3$, $y(2) = 8/3$ ft, $\dot{y}(2) = 2$ ft/sec
 25. $y = x^2$

SECTION 2.2

1. e^{3x}, e^{-3x} 3. e^x, e^{-2x} 5. 1, e^x 7. e^{ix}, e^{-ix}
 9. $e^{-(2+i)x}, e^{-(2-i)x}$ 11. Use Fundamental Theorem 1 in Sec. 2.1.
 13. $y'' + 3y' + 2y = 0$ 15. $y'' - 7y' + 12y = 0$ 17. $y'' + 9y = 0$
 19. $y'' + 6y' + 25y = 0$

SECTION 2.3

- | | | | |
|--------------------------------------|---|------------------------------------|--------------|
| 1. Independent | 3. Independent | 5. Dependent | 7. Dependent |
| 9. Independent | 13. $y = c_1 e^x + c_2 e^{-x}$ | 15. $y = c_1 e^{4x} + c_2 e^{-3x}$ | |
| 17. $y = c_1 e^{2ix} + c_2 e^{-2ix}$ | 19. $y = c_1 e^{(1+i)x} + c_2 e^{(1-i)x}$ | | |
| 21. $y'' - 3y' + 2y = 0$ | 23. $y'' + 6y' + 10y = 0$ | 25. $y'' + 4y' = 0$ | |

SECTION 2.4

- | | |
|---|---|
| 7. $A \cos \pi x + B \sin \pi x$ | 9. $e^{3x}(A \cos 4x + B \sin 4x)$ |
| 11. $\lambda^2 + a\lambda + b = (\lambda - \lambda_1)(\lambda - \lambda_2) = (\lambda - p - iq)(\lambda - p + iq) = \lambda^2 - 2p\lambda + p^2 + q^2$,
hence $a = -2p$ and $b = p^2 + q^2$. | |
| 13. $y'' + 2y' + 2y = 0$ | 15. $y'' + 6y' + 12y = 0$ |
| 17. $y = A \cos 0.5x + B \sin 0.5x$ | 21. $y = c_1 e^{-4x} + c_2 e^{-2x}$ |
| 19. $y = e^{-0.1x}(A \cos 0.5x + B \sin 0.5x)$ | 25. $y = A \cos \sqrt{5}x + B \sin \sqrt{5}x$ |
| 23. $y = e^{-ax}(A \cos x + B \sin x)$ | 29. $y = e^x \cos x$ |
| 27. $y = e^{-2.5x}(A \cos 2.5x + B \sin 2.5x)$ | 33. $y = 2 \cos \pi x - \sin \pi x$ |
| 31. $y = 2e^{-2x} \sin 3x$ | 37. $y = e^{-2x}(\cos x - \sin x)$ |
| 35. $y = 2e^{2x}$ | |
| 39. $y = e^{2x} - 2e^x$ | |

SECTION 2.5

- | | | |
|--------------------------------|----------------------------------|---------------------------------|
| 9. $y = (c_1 x + c_2)e^{0.1x}$ | 11. $y = (c_1 x + c_2)e^{-1.5x}$ | 13. $y = (c_1 x + c_2)e^{2.5x}$ |
| 15. $y = 3xe^x$ | 17. $y = (2x + 2)e^{3x}$ | 19. $y = -2e^{-0.5x}$ |

SECTION 2.6

1. (a) $y = \cos t - \sin t = \sqrt{2} \cos(t + \pi/4)$, (b) $y = \cos t$, (c) $y = \sqrt{2} \cos(t - \pi/4)$;
 $y(\pi) = -1$, $y(2\pi) = 1$; $\dot{y}(\pi) = 1, 0, -1$; $\dot{y}(2\pi) = -1, 0, 1$
3. $y = \cos \omega_0 t + (1/\omega_0) \sin \omega_0 t = \sqrt{1 + 1/\omega_0^2} \cos(\omega_0 t - \arctan(1/\omega_0))$

5. No

7. $y = \frac{1}{2\beta} \{[(\alpha + \beta)y_0 + v_0]e^{-(\alpha-\beta)t} + [(\beta - \alpha)y_0 - v_0]e^{-(\alpha+\beta)t}\}$

9. From (9) and $\dot{y} = 0$ we have $-c_1/c_2 = e^{-2\beta t}$. Since the exponential function is positive, the statement follows.

11. It decreases as c increases.

13. $y = e^{-t} \cos t$, $\dot{y} = e^{-t} (\cos t + \sin t) = \sqrt{2} e^{-t} \cos(t - \pi/4)$,
 $\ddot{y} = e^{-t} (\cos t + 2\sin t) = \sqrt{5} e^{-t} \cos(t - 1.11)$

15. From (11) and $\dot{y} = 0$ we obtain $\tan(\omega^* t - \delta) = -\alpha/\omega^* = \text{const}$, and consecutive solutions of this equation have the constant distance π/ω^* .

17. The positive solutions of $\tan t = -1$, that is, $3\pi/4$ (min.), $7\pi/4$ (max.), etc., $\Delta = 2\pi$.

19. $\dot{y} = 0$ yields $t = t_0 = 1/\alpha - c_2/c_1$, and $t_0 > 0$ if $\alpha < c_1/c_2$ or (cf. Prob. 18) $\alpha < (v_0 + \alpha y_0)/y_0$, that is, $v_0/y_0 > 0$. Hence, a maximum or minimum arises if y_0 and v_0 have the same sign.

SECTION 2.7

1. Substitution of $y_2 = uy_1$, $y_2' = uy_1' + u'y_1$, $y_2'' = uy_1'' + 2u'y_1' + u''y_1$ into (1) gives
 $(x^2 y_1'' + ax y_1' + by_1)u + (2x^2 y_1' + ax y_1)u' + x^2 u''y_1 = 0$.

Since y_1 satisfies (1), the factor of u is zero. Since $y_1' = mx^{m-1}$ and $a = 1 - 2m$ in the critical case, the remaining expression is $u' + xu'' = 0$. Thus, $u' = 1/x$, $u = \ln x$, $y_2 = x^m \ln x$.

5. $y = c_1x + c_2x^3$

11. $y = c_1 + c_2 \ln x$

17. $y = x - x^3$

19. $y + (a-1)y' + by = 0$, $x = e^t$, $t = \ln x$; by the chain rule, $y = y'x = y'e^t$, $y = y''x^2 + y'x$, $y''x^2 + y'x + (a-1)y'e^t + by = 0$

7. $y = (c_1 + c_2 \ln x)x$

13. $y = c_1x^{-2} + c_2x^{-\frac{1}{2}}$

9. $y = c_1 + c_2x^4$

15. $y = (1 - \ln x)x^2$

SECTION 2.8

1. $(m_3 - m_1)x^{m_1+m_2-1}$

3. x^{2m-1}

5. $e^{2\lambda x}$

9. $y' + y = 0$

11. $y'' - 2y' + y = 0$

13. $xy'' - y' = 0$

15. $x^2y'' - 3xy' + 4y = 0$

SECTION 2.9

5. Dependent since $\cosh^2 x - \sinh^2 x = 1$

7. Dependent

9. Dependent

11. Independent

SECTION 2.10

3. If $\lambda = k$ is a triple root, then (1) becomes (1*) $y''' - 3ky'' + 3k^2y' - k^3y = 0$, and $y_1 = e^{kx}$, $y_1' = ky_1$. Hence $y_3 = uy_1$, $y_2' = (u' + ku)y_1$, $y_2'' = (u'' + 2ku' + k^2u)y_1$, $y_3''' = (u''' + 3ku'' + 3k^2u' + k^3u)y_1$. Substitution in (1*) shows that everything cancels, except for $u'''y_1$. Hence $u''' = 0$, which has the independent solutions 1, x , x^2 .

5. $y = c_1 + c_2 \cos x + c_3 \sin x$

7. $y = c_1e^x + c_2e^{-x} + c_3e^{2x} + c_4e^{-2x}$

9. $y = A \cos x + B \sin x + C \cos 2x + K \sin 2x$

11. $y = c_1x^{-1} + c_2x + c_3x^2$

17. $y = c_1x^{-1} + c_2x + c_3x^3$

19. $y = (c_1 + c_2e^x + c_3e^{-x})x$

SECTION 2.11

9. $y = A \cos x + B \sin x + e^x$

11. $y = c_1e^x + c_2e^{2x} + x^2$

13. $y = A \cos 3x + B \sin 3x + 2x$

15. $y = A \cos x + B \sin x - x \cos x$

SECTION 2.12

1. $5x + 7.5$

3. $x^2 + x - 2$

5. $1.5x^4 - 5x^3 + 9.5x^2 - 11x + 6$

7. $-5 \cos 2x + \sin 2x$

9. $-x \cos x$

11. $2x^3 - 3x^2 + 15x - 8$

13. $A \cos x + B \sin x - x^2 - x + 2$

15. $A \cos 2x + B \sin 2x + 0.2e^{-x}$

17. $A \cos x + B \sin x - 0.5x \cos x$

19. $c_1e^{3x} + c_2e^x + 0.5xe^{3x}$

21. $ce^{-x} + x^4 - 4x^3 + 12x^2 - 24x + 24$

23. $ce^{-4x} + (4 \cos nx + \pi \sin nx)/(16 + \pi^2)$

25. $3e^x + 2x^2$

27. $e^{-x} - e^{2x} + xe^{2x}$

29. $e^{-2x} \cos 2x + \sin x$

SECTION 2.13

1. $y = A \cos t + B \sin t - \cos 2t$

3. $y = A \cos 4t + B \sin 4t + (\cos t - \sin t)/15$

5. $y = \cos 5t + \sin t$

7. $y = (v_0 - (\omega_0^2 - \omega^2)^{-1}) \cos \omega_0 t + (v_0/\omega_0) \sin \omega_0 t + (\omega_0^2 - \omega^2)^{-1} \cos \omega t$

9. No 13. $\sin t - 3 \cos t$ 15. $6 \sin 3t - 7 \cos 3t$

17. $y = A \cos \omega t + B \sin \omega t$ where $A = -b\omega c/D$, $B = b(m(\omega_0^2 - \omega^2)/D)$,

$D = m^2(\omega_0^2 - \omega^2)^2 + \omega^2c^2$, $\omega_0 = \sqrt{k/m}$

29. $e^{-t/4}(A \cos(\sqrt{31}/4)t + B \sin(\sqrt{31}/4)t) + 4 \cos t + 2 \sin t$

21. $(c_1 + c_2 t)e^{-t} - 6 \cos 3t - 8 \sin 3t$

23. $y = -0.1 \cos 5t$

25. $y = e^{-t} \sin t - 0.5 \sin 2t$

SECTION 2.14

1. $I = \cos t + 2 \sin t$

3. $I = 10(\cos 5t + \sin 5t)$

5. $I = e^{-2t}(A \cos 4t + B \sin 4t) + 10 \cos 4t + 40 \sin 4t$

9. $I = 60 \sin 2t - 30 \sin t$

7. Use (3') in Sec. 1.9.

13. $I = e^{-3t}(3 \sin 4t - 4 \cos 4t) + 4 \cos 5t$

11. $I = 5e^{-2t} \sin t$

15. $I = \begin{cases} I_1 \equiv A_1 \cos t + B_1 \sin t + 1 & \text{when } 0 < t < a, \\ I_2 \equiv A_2 \cos t + B_2 \sin t & \text{when } t > a. \end{cases}$

$I(0) = 0$, $q(0) = 0$, $E(0) = 0$ yields $I(0) = 0$, $A_1 = -1$, $B_1 = 0$, and $I_1 = 1 - \cos t$. From this, $I_1(a) = 1 - \cos a$, $I_1'(a) = \sin a$. Thus $I_2(a) = I_1(a) = 1 - \cos a$, $I_2'(a) = I_1'(a) - a = -a + \sin a$, that is, $A_2 \cos a + B_2 \sin a = 1 - \cos a$, $-A_2 \sin a + B_2 \cos a = -a + \sin a$. Hence, $A_2 = \cos a + a \sin a - 1$, $B_2 = \sin a - a \cos a$, and

1 - $\cos t$ when $0 < t < a$,

$\cos(t-a) - a \sin(t-a) - \cos t$ when $t > a$.

17. $I = \begin{cases} \frac{1}{2}(e^{-t} - \cos t + \sin t) & \text{when } 0 < t < \pi \\ -\frac{1}{2}(1 + e^{-\pi}) \cos t + \frac{1}{2}(3 - e^{-\pi}) \sin t & \text{when } t > \pi \end{cases}$

21. $I = \begin{cases} 1 - e^{-t} & \text{when } 0 < t < a \\ (e^a - 1)e^{-t} & \text{when } t > a \end{cases}$ 23. $I = \begin{cases} 1 & \text{when } 0 < t < a \\ -ae^{a-t} & \text{when } t > a \end{cases}$

SECTION 2.15

1. $I_p = \operatorname{Re}[(-0.1 - 0.2i)(\cos 2t + i \sin 2t)] = -0.1 \cos 2t + 0.2 \sin 2t$

3. $I_p = (360 \sin 3t - 300 \cos 3t)/61$ 5. $I_p = (50 \sin 4t - 110 \cos 4t)/73$

7. $y_p = 2 \sin 4t$ 9. $y_p = 0.1008 \sin 10t - 0.2006 \cos 10t$

SECTION 2.16

1. $y = A \cos x + B \sin x - x \cos x + \sin x \ln(\sin x)$

3. $y = (c_1 + c_2 x + x \ln x - x) e^{2x}$ 5. $y = (c_1 + c_2 x - \cos x) e^{-x}$

7. $y = (c_1 + c_2 x + \frac{4}{3}x^{\frac{3}{2}}) e^x$ 9. $y = e^{-x}(A \cos x + B \sin x - 0.5 \cos 2x / \cos x)$

11. $y = c_1 x^{-1} + c_2 x - 4$. Note that in (8), $r = 4/x^2$ (not 4), because the given equation must be divided by x^2 in order that it assumes the form (1) from which (8) was derived.

13. $y = c_1 x + c_2 x^2 + x^4/6$

15. $y = c_1 x + c_2 x^2 + \frac{1}{12}x^2$

17. $y = c_1 x^3 + c_2 x^3 + \frac{1}{4}x^4$

19. $y = c_1 + c_2 x^2 + (2x - 2)e^x$

SECTION 3.2

1. 3

3. 1

5. ∞

7. 1

9. ∞

11. $\frac{1}{4}$

13. 1

15. 3

17. $c_0 = 0$, $c_1 = 2$, c_n arbitrary, $c_m = c_{m-1}/(m-2)$,

$y = 2x + c_2 x^2 \left(1 + x + \frac{x^2}{2!} + \dots\right) = c_2 x^2 e^x + 2x$

19. $y = c_1(x + x^2)$

21. $y = c_0 + c_1 x + (\frac{3}{2}c_1 - c_0)x^2 + (\frac{7}{8}c_1 - c_0)x^3 + \dots$. Setting $c_0 = A + B$ and $c_1 = A + 2B$ we obtain $y = Ae^x + Be^{\frac{3}{2}x}$. This illustrates the fact that even if the solution of an equation is a known function, the power series method may not yield it immediately in the usual form.

23. $y = c_1 x + c_0(1 - x^2 - \frac{1}{3}x^4 - \frac{1}{3}x^6 - \frac{1}{7}x^8 - \dots)$. (This is a particular case of Legendre's equation ($n = 1$) which will be considered in the next section.)

25. $y = c_0 \left(1 - \frac{1}{2!} x^2 - \frac{1}{4!} x^4 - \frac{3}{6!} x^6 - \frac{3 \cdot 5}{8!} x^8 - \dots \right) + c_1 x$

27. $(t+1)_t - y = t+1, y = c_0(1+t) + (1+t) \left(t - \frac{t^2}{2} + \dots \right) = c_0 x + x \ln x$

29. $y = c_0 \left(1 + \frac{t^2}{2!} + \dots \right) + c_1 \left(t + \frac{t^3}{3!} + \dots \right) = A \cosh x + B \sinh x$

where $A = c_0 \cosh 1 - c_1 \sinh 1, B = c_1 \cosh 1 - c_0 \sinh 1$

SECTION 3.3

1. $P_6(x) = (231x^6 - 315x^4 + 105x^2 - 5)/16$

5. $y = c_0(1 - 3x^2) + c_1 \left(x - \frac{2}{3}x^3 - \frac{1}{5}x^5 - \frac{4}{5 \cdot 7}x^7 - \dots \right)$

11. Set $(x^2 - 1)^n = v(x)$. Since $v, v', v', \dots, v^{(n-1)}$ are zero at $x = \pm 1$ and $v^{(2n)} = (2n)!$, formula (13) yields

$$\begin{aligned} (2^n n!)^2 \int_{-1}^1 P_n^2 dx &= \int_{-1}^1 v^{(n)} v^{(n)} dx = [v^{(n-1)} v^{(n)}] \Big|_{-1}^1 - \int_{-1}^1 v^{(n-1)} v^{(n+1)} dx = \dots \\ &= (-1)^n \int_{-1}^1 vv^{(2n)} dx = (-1)^n (2n)! \int_{-1}^1 (x^2 - 1)^n dx = 2(2n)! \int_0^1 (1 - x^2)^n dx \\ &= 2(2n)! \int_0^{\pi/2} \sin^{2n+1} \beta d\beta = 2(2n)! \frac{2 \cdot 4 \cdots (2n)}{1 \cdot 3 \cdot 5 \cdots (2n+1)} = 2(2n)! \frac{2^n n!}{(2n+1)!/2 \cdot 4 \cdots (2n)} \end{aligned}$$

13. Multiply (15) by P_m , integrate from -1 to 1 , and use (12) and (14).

15. $8P_1 + 4P_3$

17. $x^2 = (P_0 + 2P_2)/3, x^3 = (3P_1 + 2P_3)/5$

SECTION 3.4

1. $y_1 = \frac{c_0}{x^2} \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots \right) = c_0 \frac{\sin x}{x^2}$

3. $y_1 = \frac{x^2}{3} - \frac{x^4}{2 \cdot 5} + \frac{x^6}{2^2 \cdot 7} - \frac{x^8}{2^3 \cdot 19} + \dots \quad y_2 = \frac{1}{x}$

5. $y_1 = x^{-2} \sin x^2, y_2 = x^{-2} \cos x^2$

9. $y_1 = 1 + \frac{x^2}{2^2} + \frac{x^4}{(2 \cdot 4)^2} + \frac{x^6}{(2 \cdot 4 \cdot 6)^2} + \dots \quad y_2 = y_1 \ln x - \frac{x^2}{4} - \frac{3x^4}{8 \cdot 16}$

11. $y_1 = x^2, y_2 = x^{-2} \quad 13. y_1 = x^{-3} \sinh 2x, y_2 = x^{-3} \cosh 2x$

15. $y_1 = e^x, y_2 = e^x \ln x$

23. y_1/y_2 is not a constant. This implies linear independence.

31. $y = C_1 F(1, \frac{1}{2}, -\frac{1}{2}; x) + C_2 x^{\frac{1}{2}} F(\frac{3}{2}, 2, \frac{5}{2}; x)$

33. $y = C_1 F(-\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}; t-1) + C_2 (t-1)^{\frac{1}{2}}$

35. $y = C_1 F(2, -2, -\frac{1}{2}; t-2) + C_2 x^{\frac{1}{2}} F(\frac{7}{2}, -\frac{1}{2}, \frac{5}{2}; t-2)$

SECTION 3.5

7. $J_{-1}(x) = -J_1(x)$

9. $u''v + u'(2v' + fv) + \dots, 2v' + fv = 0, v = e^{-\frac{1}{2} \int f dx}$

11. $y = (A \cos x + B \sin x)/\sqrt{x}$

13. $y = AJ_{1/4}(x) + BJ_{-1/4}(x)$

15. $y = AJ_{\nu}(\lambda x) + BJ_{-\nu}(\lambda x)$

17. $y = AJ_{\nu}(x^2) + BJ_{-\nu}(x^2)$

19. $y = \sqrt{x} (AJ_{1/4}(\frac{3}{4}x^{3/4}) + BJ_{-1/4}(\frac{3}{4}x^{3/4}))$

23. $J_0(x) = 1 - \frac{1}{4}x^2 + \frac{1}{64}x^4 - + \dots$; $1 - \frac{1}{4}x^2 = 0$ when $x = 2$, and $1 - \frac{1}{4}x^2 + \frac{1}{64}x^4 = 0$ when $x = \sqrt{8}$. Clearly,

$$J_0(2) = \frac{1}{(2!)^2} - \frac{1}{(3!)^2} + - \dots > 0, \quad J_0(\sqrt{8}) = - \frac{2^2}{(3!)^2} + \frac{2^4}{(4!)^2} - + \dots < 0,$$

because the terms decrease monotone in absolute value.

25. Approximate values $\frac{1}{4}\pi + k\pi = 0.785, 3.927, 7.069, 10.210, 13.352$

SECTION 3.6

1. Start from (2), with v replaced by $v - 1$, and replace J_{v-1} by using (1).

7. $J_3 = (8x^{-2} - 1)J_1 - 4x^{-1}J_0$, cf. (3) and Prob. 6.

17. $-J_0(x) + c$

21. $\int v^3 J_0 dv = \int v^2 (v J_0) dv = v^2 (v J_1) - 2 \int v^2 J_1 dv = v^3 J_1 - 2v^2 J_2 + c$

SECTION 3.7

1. $AJ_0(x) + BY_0(x)$

5. $x^{-1}[AJ_1(x) + BY_1(x)]$

9. $AJ_2(\lambda x) + BY_2(\lambda x)$

15. Use (12) in Sec. 3.5.

3. $AJ_0(\sqrt{x}) + BY_0(\sqrt{x})$

7. $x^{-2}[AJ_2(x) + BY_2(x)]$

11. $\sqrt{x}[4J_0(\sqrt{x}) + BY_0(\sqrt{x})]$

19. Use (16) in Sec. 3.5.

SECTION 4.1

3. $\frac{2}{\sqrt{3}} - \frac{2}{\sqrt{5}}$

5. $\frac{24}{3^4} + \frac{2}{\sqrt{3}}$

7. $\frac{a^3}{\sqrt{2} + 4}$

9. $\frac{1}{2\sqrt{3}} + \frac{1}{2\sqrt{2} + 8}$, cf. (10) in Sec. 0.1.

11. $\frac{\sqrt{2} + 2}{\sqrt{2} + 4 + 5}$

13. $\frac{2\sqrt{3} \cos \theta - 2m \sin \theta}{\sqrt{2} + m^2}$

15. $\frac{1}{2\sqrt{3} + 2} + \frac{1}{2\sqrt{2} + 4\sqrt{3} + 10}$

17. $\frac{3\sqrt{3}}{2\sqrt{2} - 32} + \frac{3}{2\sqrt{3}}$, cf. Prob. 18 in Sec. 0.1.

19. $F(s) = k \int_1^2 e^{-st} dt = \frac{k}{s}(e^{-s} - e^{-2s})$

21. $\frac{1}{s^2}(1 - e^{-s}) - \frac{1}{s}e^{-2s}$

31. $\frac{1}{3} \sin 3t$

33. $3 \cos t - 2 \sin t$

35. $e^{-2t} \cos \omega t$

37. $(e^{at} - e^{bt})/(a - b)$

SECTION 4.2

11. $\frac{1}{2}(e^{2t} - \frac{1}{2})$

13. $\cosh t - 1$

15. $2(1 - e^{-t}) - t$

17. $2e^{3t} - 9t^2 - 6t - 2$

19. $6 \sin \pi t + \pi^3 t^3 - 6\pi t$

23. $f' = 1$ ($0 < t < 1$), $\mathcal{L}(f') = (1 - e^{-s})/s$ (by direct integration),

$\mathcal{L}(f') = s\mathcal{L}(f) - 0 - (0 - 1)e^{-s}$. Ans. $(1 - e^{-s})/s^2 - e^{-s}/s$

25. $(e^{-s} - e^{-2s})/s^2 - e^{-2s}/s$

SECTION 4.3

3. $y = \cos 2t + \frac{1}{8} \sin 2t$

5. $y = -2 \cosh 4x$

9. $y = -e^{-t} + 2e^{3t}$

11. $y = -e^{-3t} + e^t$

15. $y = (1 - t)e^{-t}$

17. $y = e^{0.5t} \cos 3t$

13. $y = e^{2t} (\cos t + \sin t)$

19. $y = 3e^{t/3}$

SECTION 4.5

1. $2 + e^t$

3. $1 + 3 \sinh 3t$

5. $e^t + 2e^{2t} - 3e^{-2t}$

7. $4 \cos 4t + \sin 4t$

9. $1 - \cos 2t - \sin 2t$

11. $(1 - t)e^{-t}$

13. $(1 + t^2 - 2t^3)e^{-2t}$

19. $te^{2t}(\cos t + \sin t)$

23. $y = 2e^{2t} + e^{-t}$

29. $y = e^{-t} \cos t - 2 \cos 2t - \sin 2t$

33. $y = te^{2t} + \cos 2t$

43. $y_1 = \cos t, y_2 = \sin t$

15. $2e^{2t} + te^t$

21. $e^t(\cos t + t \sin t)$

25. $y = e^{-t} \sin t$

35. $y = 2 \cos 3t + t \sin 3t$

45. $y_1 = \sin t + \cos 2t, y_2 = \sin t - \cos 2t$

17. $te^{-t} \cos t$

27. $y = e^{-2t} \sin t + 1$

31. $y = e^{3t} + e^t$

SECTION 4.6

1. $\frac{1}{(s-1)^2}$

3. $\frac{6s}{(s^2+9)^2}$

5. $\frac{2s(s^2-3)}{(s^2+1)^3}$

7. $\frac{s^2+2s+5}{(s^2+2s-3)^2}$

9. $\frac{2(s+2)}{(s^2+4s+5)^2}$

13. $t \sin t$

15. $t^{-1}e^{-t} \sin t$

SECTION 4.7

1. $\frac{k}{s}(1 - e^{-as})$

5. $1/s(1 - e^{-s})$

9. $s^{-1}[3(1 - e^{-3s})^{-1} - (1 - e^{-s})^{-1} - 1]$

15. $f = -1$ when $1 < t < 3$ and 0 otherwise

17. $f = 0$ when $t < 1, f = t - 1$ when $t > 1$, cf. Theorem 3 in Sec. 4.2.

19. $(1 - e^{-s})/s^2$

3. $\frac{1}{s}(2 - e^{-s} - e^{-2s} - e^{-3s} + e^{-4s})$

7. $k(e^{-as} - 1)/s(e^{-as} + 1)$

13. $f = 1$ when $t > 2$ and 0 otherwise

21. $(1 - 2e^{-s} + 2e^{-3s} - e^{-4s})/s^2$.

25. $V_0(1 - e^{-Rt/L})/R$

SECTION 4.8

1. e^{-s}/s^2

7. $e^{a(s-k)}/(s-k)$

13. $Kw(1 + e^{-\pi s/\omega})/(s^2 + \omega^2)$

17. $Ks^{-2}(1 - e^{-as}(1 + as))$

23. $-u_1(t)\pi^{-1} \sin \pi t$

3. $e^{-s}(s^{-2} + s^{-1})$

9. $e^{-\pi(s+2)}/(s+2)$

13. $we^{2\pi s/\omega}/(s^2 + \omega^2)$

17. $Ks(1 + e^{-\pi s/\omega})/(s^2 + \omega^2)$

21. $\frac{1}{2}(t-1)^2 u_1(t)$

25. $\frac{1}{2} \sin 2t$ when $0 < t < \pi$ and 0 otherwise

27. $i' + \int_0^t l(\tau) d\tau = i[1 - u_a(t)] = t - (t-a)u_a(t) - au_a(t).$ Ans. $i = 1 - \cos t$ when $0 < t < a, i = \cos(t-a) - a \sin(t-a) - \cos t$ when $t > a.$

29. $i = \frac{1}{2}(e^{-t} - \cos t + \sin t)$ when $0 < t < \pi,$

$i = -\frac{1}{2}(1 + e^{-\pi}) \cos t + \frac{1}{2}(3 - e^{-\pi}) \sin t$ when $t > \pi$

SECTION 4.9

3. $\frac{\omega}{(s^2 + \omega^2)(e^{\pi s/\omega} - 1)}$

5. $\frac{\omega}{s^2 + \omega^2} \coth \frac{\pi s}{2\omega}$

7. $\frac{1}{s^2 + \omega^2} \left(s + \frac{\omega}{\sinh \pi s/2\omega} \right)$

9. $\frac{e^{2(1-s)\pi} - 1}{(1-s)(1-e^{-2\pi s})}$

11. Set $\omega = \frac{1}{2}$ in Prob. 5 and multiply by $K.$

13. $\left[\frac{\pi}{s} e^{-\pi s} (e^{-\pi s} - 1) + \frac{1}{s^2} (e^{-\pi s} - 1)^2 \right] / (1 - e^{-2\pi s})$

15. $\left[\frac{1}{s^2} (1 - e^{-\pi s}) - \frac{\pi}{s} e^{-\pi s} \right] / (1 - e^{-2\pi s})$

19. $Ri + \frac{1}{C} \int i dt = v(t)$, $RI + \frac{I}{Cs} = \frac{1 - e^{-s}}{s^2} - \frac{e^{-s}}{s}$, $i = C(1 - e^{-t/RC})$ when $0 < t < 1$,
 $i = ((C - R^{-1})e^{1/RC} - C)e^{-t/RC}$ when $t > 1$

21. $Ri + C^{-1}q = v(t)$, and since the charge q on the capacitor cannot change abruptly, the current i must have a jump when v has a jump.

23. $i' + \frac{R}{L}i = \frac{V_0}{L}[u_0(t) - u_a(t) + \dots]$, $I = \frac{V_0}{L} \frac{1}{s(s+k)}[1 - e^{-as} + e^{-2as} - \dots]$,

$$k = \frac{R}{L}, \quad \mathcal{L}^{-1}\left\{\frac{1}{s(s+k)}\right\} = 1 - e^{-kt}, \quad \mathcal{L}^{-1}\left\{\frac{e^{-as}}{s(s+k)}\right\} = (1 - e^{-k(t-a)})u_a(t), \text{ etc.,}$$

$$i = \frac{V_0}{R}(1 - e^{-kt}) \quad (t < a), \quad i = -\frac{V_0}{R}e^{-kt}(1 - e^{ka}) \quad (a < t < 2a), \text{ etc.}$$

In general, by summing the finite geometric progression,

$$i = \frac{V_0}{R} \left[\delta - \frac{e^{-kt}}{1 + e^{ka}} - (-1)^n \frac{e^{-k[t-(n+1)a]}}{1 + e^{ka}} \right], \quad \delta = \begin{cases} 0 & (n \text{ odd}), \\ 1 & (n \text{ even}), \end{cases} \quad na < t < (n+1)a.$$

The second term on the right dies out. If we set $t - na = \tau$, the remaining expression becomes $i = V_0 R^{-1} (1 - e^{-k\tau} (1 + e^{-ka})^{-1})$ ($n = 0, 2, \dots$), $i = V_0 R^{-1} e^{-k\tau} (1 + e^{-ka})^{-1}$ ($n = 1, 3, \dots$), $0 < \tau < a$.

27. $i = e^{-\alpha t} \left(i_0 \cos \omega^* t + \frac{V_0 + (\alpha L - R)i_0}{\omega^* L} \sin \omega^* t \right)$

29. $i = R^{-1}t + R^{-2}L(e^{-tR/L} - 1) \quad (0 < t < 1),$
 $i = (R^{-1}e^{R/L}(1 - R^{-1}L) + R^{-2}L)e^{-tR/L} \quad (t > 1)$

31. $i(t) = \frac{1}{R} \left(\frac{e^{-kt}}{1 - e^{-k}} + t - \frac{L}{R} \right)$ when $0 < t < 1$, and $i(t+1) = i(t)$, $k = \frac{R}{L}$

33. $i_1 = V_0 \left[e^{-\alpha t} \left(-\frac{1}{R_1} \cos \omega^* t + \frac{1}{\omega^*} \left(\frac{L_2}{A} - \frac{\alpha}{R_1} \right) \sin \omega^* t \right) + \frac{1}{R_1} \right]$

35. $L_2 i_2' + R_2 i_2 = M i_1'$ where $i_1(0) = V_0/R_1$, $i_1 = 0$ when $t > 0$ and thus

$$L_2 s I_2 + R_2 I_2 = M \mathcal{L}(i_1') = M[s \mathcal{L}(i_1) - i_1(0)] = -\frac{MV_0}{R_1}, \quad i_2 = -\frac{MV_0}{R_1 L_2} e^{-R_2 t / L_2}$$

37. $y = \frac{P}{K} \left[1 + \frac{1}{2\beta} (\alpha - \beta) e^{-(\alpha+\beta)t} - \frac{1}{2\beta} (\alpha + \beta) e^{-(\alpha+\beta)t} \right]$ where $\beta^2 = -\omega^*{}^2$

39. Let ϕ_1 and ϕ_2 be the angular displacements of the wheels such that at $t = 0$, $\dot{\phi}_1 = \dot{\phi}_2 = 0$, $\phi_1' = \dot{\phi}_2' = \omega$. Let c be the stiffness of the shaft (couple per radian relative twist of the wheels). Then at $t > 0$,

$$M_1 \phi_1'' - c(\phi_2 - \phi_1) = -P, \quad M_2 \phi_2'' + c(\phi_2 - \phi_1) = 0.$$

Let $\Phi_1 = \mathcal{L}(\phi_1)$, $\Phi_2 = \mathcal{L}(\phi_2)$. Then the subsidiary equations are

$$(M_1 s^2 + c)\Phi_1 - c\Phi_2 = M_1 \omega - \frac{P}{s}, \quad -c\Phi_1 + (M_2 s^2 + c)\Phi_2 = M_2 \omega.$$

Elimination of Φ_1 yields

$$\Phi_2 = \frac{\omega}{s^2} - \frac{cP}{s^3 [M_1 M_2 s^2 + c(M_1 + M_2)]}, \text{ and } V = \mathcal{L}(v) = s\Phi_2.$$

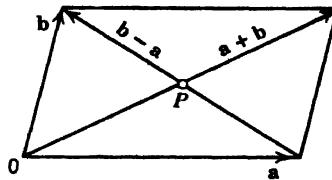
$$\text{Ans. } v = \omega - \frac{Pt}{M_1 + M_2} + \frac{P}{k(M_1 + M_2)} \sin kt, \quad k^2 = c \left(\frac{1}{M_1} + \frac{1}{M_2} \right)$$

SECTION 5.2

1. $2, 0, 1, |\mathbf{v}| = \sqrt{5}$ 3. $1, 1, 1, |\mathbf{v}| = \sqrt{3}$ 5. $2, 3, 1, |\mathbf{v}| = \sqrt{14}$
 7. $0, -1, 0, |\mathbf{v}| = 1$ 9. $0, 0, 0, |\mathbf{v}| = 0$ 11. $Q, (7, 4, 5)$
 13. $Q: (-\frac{3}{2}, \frac{5}{2}, -1)$ 15. $Q: (0, 0, 0)$ 17. $Q: (3, 8, -9)$
 19. $Q: (-3, -2, -2)$

SECTION 5.3

1. $-2\mathbf{i} - 4\mathbf{j} + 6\mathbf{k}, \frac{1}{2}\mathbf{i} + \mathbf{j} - \frac{3}{2}\mathbf{k}, 3\mathbf{i} + 6\mathbf{j} - 9\mathbf{k}$
 3. $-\mathbf{i} + \mathbf{j} - 7\mathbf{k}, \mathbf{i} - \mathbf{j} + 7\mathbf{k}, 2\mathbf{i} + 6\mathbf{j} + 4\mathbf{k}$ 5. $\sqrt{59}, \sqrt{14} - 5$
 7. $-\mathbf{i} + 4\mathbf{j} - 17\mathbf{k}, \mathbf{i} - 4\mathbf{j} + 17\mathbf{k}$ 9. $\sqrt{19} + 5, \sqrt{14} + 6$
 11. $\mathbf{i} - \mathbf{j}$ 13. $\mathbf{a} + \mathbf{b} + \mathbf{c}, \mathbf{a} - \mathbf{b} + \mathbf{c}, \mathbf{a} + \mathbf{b} - \mathbf{c}, \mathbf{a} - \mathbf{b} - \mathbf{c}$
 15. The vector \mathbf{r} from 0 to P is of the form $\mathbf{r} = k(\mathbf{a} + \mathbf{b})$. Also, $\mathbf{r} = \mathbf{a} + l(\mathbf{b} - \mathbf{a})$. Thus, $k\mathbf{a} + k\mathbf{b} = (1 - l)\mathbf{a} + l\mathbf{b}$. From this, $k = 1 - l$, $k = l$, hence $k = \frac{1}{2}$, $l = \frac{1}{2}$, and the proof is complete.



Problem 15

SECTION 5.4

1. $5, 5, 0$ 3. $-11, -11$ 5. $\sqrt{41}, \sqrt{43}$ 7. $\sqrt{18}, \sqrt{14} + \sqrt{26}$
 9. 27 11. -2 13. -4
 15. The force is orthogonal to the segment
 17. $(5\mathbf{i} + 2\mathbf{j})/\sqrt{29}$ 19. No 21. $23/26$ 23. $25/2\sqrt{195}$
 27. The angle β of the triangle at B is the angle between the vectors $\mathbf{c} = -2\mathbf{i}$ from B to A and $\mathbf{a} = -2\mathbf{i} + 2\mathbf{j}$ from B to C . Therefore $\cos \beta = \mathbf{a} \cdot \mathbf{c}/|\mathbf{a}| |\mathbf{c}| = 1/\sqrt{2}$, $\beta = \pi/4$, etc.

SECTION 5.6

1. $6\mathbf{i} - 3\mathbf{j}, -6\mathbf{i} + 3\mathbf{j}$ 3. $3\mathbf{i} + 3\mathbf{j} + 3\mathbf{k}, 1$ 5. $-12\mathbf{i} + 6\mathbf{j} - 6\mathbf{k}$
 7. $21\mathbf{i} + 3\mathbf{j} + 9\mathbf{k}$ 9. $3\mathbf{i} - 6\mathbf{j} - 3\mathbf{k}$ 11. 0
 13. $3\mathbf{i} - 6\mathbf{j}$ 15. $6\mathbf{i} - 3\mathbf{j}$ 17. $6\sqrt{14}$ 19. $\sqrt{141}$
 21. $\sqrt{14}$ 23. $\sqrt{86}$ 25. $\sqrt{8}$ 27. 3
 29. 2 31. 2 33. $\frac{5}{2}$ 35. $\frac{1}{2}\sqrt{27}$
 37. Parallel 39. No 41. Parallel 43. $\mathbf{k}, -\mathbf{k}$
 45. $\mathbf{i}, -\mathbf{i}$ 47. $\pm(\mathbf{i} + \mathbf{j} - 2\mathbf{k})/\sqrt{6}$
 49. $\pm(2\mathbf{i} + 2\mathbf{j} + \mathbf{k})/3$ 51. $\pm(7\mathbf{i} - \mathbf{j} - 3\mathbf{k})/\sqrt{59}$
 53. $\mathbf{i} - \mathbf{j}$ 55. \mathbf{k} 57. $\mathbf{i} - 2\mathbf{j} + \mathbf{k}$

SECTION 5.7

1. 1 3. -1 5. 3 7. -12
 9. $\leftarrow 3$ 11. Independent 13. Independent 15. Dependent

17. Independent 19. 3
25. $\frac{7}{3}$ 27. $\frac{2}{3}$

21. 2 23. 20 *
29. $\frac{5}{3}$

SECTION 5.8

1. $6\mathbf{i} + 6\mathbf{j} - 7\mathbf{k}, 9$
7. -15
13. $21\mathbf{i} + 21\mathbf{j} - 7\mathbf{k}$
3. $6\mathbf{i} + 3\mathbf{j} - 3\mathbf{k}, 5\mathbf{i} + 5\mathbf{j}$
9. $-11\mathbf{i} + 7\mathbf{j} - 12\mathbf{k}$
15. $-17\mathbf{i} + 26\mathbf{j} + 27\mathbf{k}$
5. $10\mathbf{i} - 7\mathbf{j} - 4\mathbf{k}, \sqrt{165}$
11. $12\mathbf{i} + 12\mathbf{j} - 14\mathbf{k}$
17. L is parallel to $P.$

SECTION 5.9

13. $T = 3y - 4y^3$
17. Concentric spheres
21. 0, -1, 0
25. Concentric circles, parabolas
37. $x^2 + y^2 = \text{const}$, straight lines $2ry/(x^2 - y^2) = \text{const}$
39. Circles, straight lines
15. Parallel planes
19. Paraboloids of revolution
23. Paraboloid of revolution
33. $6x^4\mathbf{i} - 6x^2\mathbf{j}$

SECTION 5.10

7. $\mathbf{u}' = \mathbf{i} + 2t\mathbf{j}, \mathbf{u}'' = 2\mathbf{j}$
9. $\mathbf{u}' = -\sin t\mathbf{i} + 4\cos t\mathbf{j}, \mathbf{u}'' = -\cos t\mathbf{i} - 4\sin t\mathbf{j}$
11. $\mathbf{u}' = -2\sin t\mathbf{i} + 2\cos t\mathbf{j} + \mathbf{k}, \mathbf{u}'' = -2\cos t\mathbf{i} - 2\sin t\mathbf{j}$
13. $-10t^4\mathbf{i} - 2t\mathbf{j} + 4t^3\mathbf{k}$
15. $-(4t^3 + 5t^4)\mathbf{i} + (4t^3 + 14t^6)\mathbf{j} - (12t^5 - 2t)\mathbf{k}$
17. $\mathbf{l} + 3t^2 + 12t^3$
21. $\mathbf{z}\mathbf{j} + y\mathbf{k}, z\mathbf{i} + x\mathbf{k}, y\mathbf{i} + x\mathbf{j}$
25. $(\mathbf{u} \cdot \mathbf{v})'' = \mathbf{u}'' \cdot \mathbf{v} + 2\mathbf{u}' \cdot \mathbf{v}' + \mathbf{u} \cdot \mathbf{v}'', (\mathbf{u} \times \mathbf{v})'' = \mathbf{u}'' \times \mathbf{v} + 2\mathbf{u}' \times \mathbf{v}' + \mathbf{u} \times \mathbf{v}''$

SECTION 5.11

1. $\mathbf{r}(t) = t\mathbf{i} + t\mathbf{j}$
5. $\mathbf{r}(t) = t\mathbf{i} + t\mathbf{j} + t\mathbf{k}$
9. $\mathbf{r}(t) = t\mathbf{i} + t\mathbf{j}$
13. $\mathbf{r}(t) = \cos t\mathbf{i} + \sin t\mathbf{j}$
17. $\mathbf{r}(t) = (-1 + \cos t)\mathbf{i} + 2 \sin t\mathbf{j}$
19. Circle in the xy -plane, sine curve in the yz -plane, cosine curve in the xz -plane
3. $\mathbf{r}(t) = 2\mathbf{i} + (1 + 2t)\mathbf{j} + t\mathbf{k}$
7. $\mathbf{r}(t) = (1 - t)\mathbf{i} + (5 - 3t)\mathbf{j} + (3 - 4t)\mathbf{k}$
11. $\mathbf{r}(t) = t\mathbf{i} - t\mathbf{j} + t\mathbf{k}$
15. $\mathbf{r}(t) = t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k}$

SECTION 5.12

1. $\sinh 1$
5. Start from $\mathbf{r}(t) = t\mathbf{i} + f(t)\mathbf{j}.$
3. $8(\sqrt{1000} - 1)/27$
9. $8a$

SECTION 5.13

3. $\mathbf{q}(w) = (1 + w)\mathbf{i} + (1 + 2w)\mathbf{j} + (1 + 3w)\mathbf{k}$
5. $\mathbf{q}(w) = (w + 2\sqrt{2})\mathbf{i} + (w - 2\sqrt{2})\mathbf{j}$
7. $\mathbf{r}' = \dot{\mathbf{r}}', t' = 1/\dot{s} = 1/\sqrt{\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}}, \text{ cf. (3) in Sec. 5.12, } \mathbf{r}'' = \ddot{\mathbf{r}}t'^2 + \dot{\mathbf{r}}t'',$
 $t'' = -\frac{1}{2}(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}})^{-\frac{3}{2}} 2(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}})t' = -\ddot{\mathbf{r}} \cdot \dot{\mathbf{r}}/(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}})^2.$ Thus $\mathbf{r}'' = \ddot{\mathbf{r}}/(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}) - \dot{\mathbf{r}}(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}})/(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}})^2.$ Taking $\sqrt{\mathbf{r}'' \cdot \mathbf{r}''}$ and simplifying, we obtain (5').
15. $\mathbf{r}' = \dot{\mathbf{r}}/\dot{s}, \mathbf{r}'' = \ddot{\mathbf{r}}/\dot{s}^2 + \dots, \mathbf{r}' = \ddot{\mathbf{r}}/\dot{s}^3 + \dots,$

$$: (\mathbf{r}' \mathbf{r}'' \mathbf{r}''')/\dot{s}^2 : \frac{(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}})^3}{\dot{s}^8} \frac{(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}})}{(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}})(\ddot{\mathbf{r}} \cdot \ddot{\mathbf{r}}) - (\dot{\mathbf{r}} \cdot \dot{\mathbf{r}})^2}$$

where the dots denote terms which disappear later by applying familiar rules for simplifying determinants.

SECTION 5.14

1. $\mathbf{v} = \mathbf{w} \times \mathbf{r}$, $\mathbf{w} = \omega \mathbf{k}$, $\dot{\mathbf{v}} = \mathbf{w} \times \dot{\mathbf{r}} = \mathbf{w} \times \mathbf{v}$. Since \mathbf{v} is tangent to the path, $\mathbf{w} \times \mathbf{v}$ has the direction opposite to \mathbf{r} , and $|\mathbf{a}| = |\mathbf{w}| |\mathbf{v}| = \omega^3 R$, as before.

3. $\mathbf{v} = \mathbf{i}$, $|\mathbf{v}| = 1$, $\mathbf{a} = \mathbf{0}$

5. $\mathbf{v} = (2 - 2t)\mathbf{i}$, $|\mathbf{v}| = |2 - 2t|$, $\mathbf{a} = -2\mathbf{i}$

7. Segment $-1 \leq y \leq 1$ on the y -axis, $\cos t\mathbf{i}$, $|\cos t|$, $-\sin t\mathbf{j}$

9. Circle, $-2t \sin t^2\mathbf{i} + 2t \cos t^2\mathbf{j}$, $2t$, $(-2 \sin t^2 - 4t^2 \cos t^2)\mathbf{i} + (2 \cos t^2 - 4t^2 \sin t^2)\mathbf{j}$

SECTION 5.15

1. $2t + 1/t$

3. $(g'h - gh')/h^2$

7. $-t(1 + t^2)^{-\frac{3}{2}}$

9. 2
 0

11. $e^{2u} \sin 2v$, $e^{2u} \cos 2v$

15. $r = \sqrt{x^2 + y^2}$, $\theta = \arctan(y/x)$, $\mathbf{r}_x = \mathbf{x}/r$, $\mathbf{r}_x = -y/r^2$, $\mathbf{r}_{xx} = y^2/r^3$, etc.,

$w_{xx} = x^2 r^{-2} w_{rr} - 2xyr^{-3} w_{r\theta} + y^2 r^{-4} w_{\theta\theta} + y^2 r^{-3} w_r + 2xyr^{-4} w_\theta$, etc.

SECTION 5.16

1. $2\mathbf{i} - \mathbf{j}$

3. $\cos x \cosh y \mathbf{i} + \sin x \sinh y \mathbf{j}$

5. $2(x\mathbf{i} + y\mathbf{j})/(x^2 + y^2)$

7. $y\mathbf{z}\mathbf{i} + x\mathbf{z}\mathbf{j} + xy\mathbf{k}$

9. $2x\mathbf{i} + 2y\mathbf{j} + 2z\mathbf{k}$

11. $(y\mathbf{z}\mathbf{i} + x\mathbf{z}\mathbf{j} + xy\mathbf{k})e^{xz^2}$

13. $2\mathbf{i} - 3\mathbf{j}$

15. $y\mathbf{i} + x\mathbf{j}$

17. $18x\mathbf{i} + 2y\mathbf{j}$

19. $\mathbf{i} - \mathbf{j}$

21. $3\mathbf{i} - \mathbf{j}$

23. $48\mathbf{i} + 16\mathbf{j}$

25. $\mathbf{i} + \mathbf{j} + \mathbf{k}$

27. $4\mathbf{i} + 4\mathbf{j} - \mathbf{k}$

29. $4\mathbf{i} + 4\mathbf{k}$

31. $x - y + z$

33. $\frac{1}{2}x^2 + y^2 + \frac{1}{2}z^2$

35. $\frac{1}{2} \ln(x^2 + y^2)$

37. $4, 4\sqrt{2}, 4, 0, -4, -4\sqrt{2}, -4, 0$

39. 3

41. 0

43. $8/\sqrt{5}$

45. $14^{-\frac{3}{2}}$

SECTION 5.17

3. $a_{11} = a_{22} = a_{33} = 1$, $b_1 = -1$, $b_2 = 5$, $b_3 = 2$, all others zero

5. $a_{12} = a_{21} = a_{33} = 1$, all others zero

7. $a_{12} = a_{23} = a_{31} = 1$, all others zero

SECTION 5.18

1. 2

3. $2x$

5. $yz + zx + xy$

7. $-x^3$

9. No, because (1) involves coordinates.

15. 0

19. 272

SECTION 5.19

1. $-2\mathbf{k}$

3. $2y\mathbf{i} + 2x\mathbf{j} + 2x\mathbf{k}$

5. $y\mathbf{i} + x\mathbf{j}$

7. $e^{xy^2}(x(z-y)\mathbf{i} + y(x-z)\mathbf{j} + z(y-x)\mathbf{k})$

9. $\operatorname{curl} \mathbf{v} = 0$, $\operatorname{div} \mathbf{v} = 2$, compressible, $\mathbf{r} = c_1 e^t \mathbf{i} + c_2 e^t \mathbf{j} + c_3 \mathbf{k}$

11. $\operatorname{curl} \mathbf{v} = -3y^2\mathbf{k}$, $\operatorname{div} \mathbf{v} = 0$, incompressible, $\mathbf{r} = (c_2 e^t + c_1)\mathbf{i} + c_3 \mathbf{j} + c_3 \mathbf{k}$

17. $(2yz - zx + xy)\mathbf{i} + (2zx - xy + yz)\mathbf{j} + (2xy - yz + zx)\mathbf{k}$,

$x^2 + y^2 + z^2 - yz - zx - xy$

19. Let c_1, c_2, c_3 and c_1^*, c_2^*, c_3^* be the components of $\operatorname{curl} \mathbf{v}$ with respect to right-handed Cartesian coordinates x_1, x_2, x_3 and x_1^*, x_2^*, x_3^* , respectively. Using (5'), Sec. 5.17, the

chain rule for functions of several variables, and (9), Sec. 5.17, we find

$$\begin{aligned}
 c_1 &= \frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3} = \sum_{m=1}^3 \left(a_{m3} \frac{\partial v_m}{\partial x_2} - a_{m2} \frac{\partial v_m}{\partial x_3} \right) \\
 &= \sum_{m=1}^3 \sum_{l=1}^3 \left(a_{m3} \frac{\partial v_m}{\partial x_l} \frac{\partial x_l}{\partial x_2} - a_{m2} \frac{\partial v_m}{\partial x_l} \frac{\partial x_l}{\partial x_3} \right) = \sum_{m=1}^3 \sum_{l=1}^3 (a_{m3} a_{l2} - a_{m2} a_{l3}) \frac{\partial v_m}{\partial x_l} \\
 &= (a_{33} a_{22} - a_{32} a_{23}) \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3} \right) + \dots \\
 &= (a_{33} a_{22} - a_{32} a_{23}) c_1 + (a_{13} a_{32} - a_{12} a_{33}) c_2 + (a_{23} a_{12} - a_{22} a_{13}) c_3.
 \end{aligned}$$

Using (3) in Sec. 5.17, Lagrange's identity (Sec. 5.8), and $\mathbf{k}^* \times \mathbf{j}^* = -\mathbf{i}^*$, $\mathbf{k} \times \mathbf{j} = -\mathbf{i}$, we obtain

$$a_{33} a_{22} - a_{32} a_{23} = (\mathbf{k}^* \times \mathbf{j}^*) \cdot (\mathbf{k} \times \mathbf{j}) = \mathbf{i}^* \cdot \mathbf{i} = a_{11}, \text{ etc.}$$

Hence $c_1 = a_{11} c_1^* + a_{21} c_2^* + a_{31} c_3^*$, which is of the form (5'), Sec. 5.17. In the same way we may obtain corresponding formulas for c_2 and c_3 . If the $x_1 x_2 x_3$ -coordinates are left-handed, then $\mathbf{k} \times \mathbf{j} = +\mathbf{i}$, but then there is a minus sign in front of the determinant in (1). This proves the theorem.

SECTION 6.2

- | | | | |
|--------------------------|-----------|-----------------------------------|------------|
| 1. $\frac{5}{3}\sqrt{5}$ | 3. 4π | 5. $\frac{5}{3}$ | 7. $\ln 2$ |
| 9. $-\frac{4}{5}^8$ | 11. 0 | 13. $\frac{7}{5}$ | 15. 8 |
| 17. $-\frac{2}{2}^5$ | | 19. $ W \leq 2, W = \frac{5}{6}$ | |

SECTION 6.3

- | | | | |
|---|------------------|----------------------------|--------------------------------|
| 1. $\frac{2}{3}$ | 3. $\frac{5}{6}$ | 5. $\frac{67}{120}$ | 7. $\frac{1}{6}$ |
| 9. $\ln 2$ | 11. 1 | 13. 1 | 15. $\frac{8}{3}, \frac{4}{3}$ |
| 17. $I_x = bh^3/12, I_y = b^3h/4$ | | | |
| 19. $I_x = (a+b)h^3/24, I_y = h(a^4 - b^4)/48(a-b)$ | | | |
| 21. 18 | 23. $32\pi/3$ | 25. $\pi(e^{-1} - e^{-4})$ | |

SECTION 6.4

- | | | | |
|-------------|--------------------|--------------------|-----------------------|
| 1. 0 | 3. -1 | 5. 4 | 7. $\pi(\cosh 1 - 1)$ |
| 9. 0 | 11. $\frac{1}{4}$ | 13. $\frac{16}{5}$ | 15. $\ln 2$ |
| 17. 32π | 19. $e^2 + 2e - 3$ | 21. -1 | |

SECTION 6.5

1. xy -plane, parallel straight lines
3. Cylinder of revolution $x^2 + y^2 = 1$, straight lines, circles
5. Elliptic cylinder $x^2 + 16y^2 = 16$, straight lines, ellipses
7. Cone of revolution $z = \sqrt{x^2 + y^2}$, circles, straight lines
9. $\mathbf{r} = u\mathbf{i} + v\mathbf{k}$
13. $\mathbf{r} = u\mathbf{i} + v\mathbf{j} + u^2\mathbf{k}$
17. $x^2/a^2 - y^2/b^2 - z = 0$
11. $\mathbf{r} = u\mathbf{i} + v\mathbf{j} + (1-u-v)\mathbf{k}$
15. $x^2/a^2 + y^2/b^2 + z^2/c^2 - 1 = 0$

SECTION 6.6

1. $\mathbf{r}^*(p, q) = \mathbf{r} + p\mathbf{r}_u + q\mathbf{r}_v$, where $\mathbf{r}, \mathbf{r}_u, \mathbf{r}_v$ refer to P
3. $(\mathbf{r}^* - \mathbf{r}) \cdot \operatorname{grad} f = 0$
5. $x + y - z = 1$
7. $4x + z = 4$

9. $(-z_x \mathbf{i} - z_y \mathbf{j} + k)/\sqrt{1 + z_x^2 + z_y^2}$

15. $(1 + 4u^2) du^2 + dv^2$

19. $(1 + 4u^2) du^2 + u^2 dt^2$

13. $4 du^2 + 9 dv^2$

17. $(1 + v^2) du^2 + 2uv du dv + (1 + u^2) dv^2$

21. 2π 23. $\pi(\sqrt{125} - 1)/6$

SECTION 6.7

1. 6π

7. 0

15. $2\pi^2 ab(2a^2 + 3b^2)$

3. $(3^{1/2} - 2^{1/2} + 1)/15$

9. $(5^{1/2} - 1)/2$

5. $(10^{3/2} - 1)/9$

13. $2\pi h$

17. $2\pi^2 ab(4a^2 + 4ab + 5b^2)$

SECTION 6.9

1. 1

7. $i\frac{1}{2}0$

13. Take $\mathbf{u} = xi, y\mathbf{j}, zk$, respectively, in (3), Sec. 6.8.

19. 3

25. Put $f = 1$ in (9).

3. $\pi/2$

9. $2c^5/3$

21. 0

5. $\frac{1}{8}$

11. $\pi h^3/10$

23. 12

27. Use Prob. 26.

SECTION 6.11

1. ± 1

7. $-18\pi\sqrt{2}$

15. 0

3. $\pm 3\pi/2$

9. $\frac{1}{3}$

5. Use a parametric representation of C .

11. 0

13. 0

SECTION 6.12

1. Yes

9. $u = (x^2 - y^2 - z^2)/2$

15. $\frac{1}{2}$

3. No

11. $u = x + yz$

17. 5

5. Yes

7. Yes

13. $u = \sin x - y^2 z$

19. $e^2 - 5$

SECTION 7.1

1. $\begin{pmatrix} 0 & 1 \\ 4 & 7 \end{pmatrix}$

3. $\begin{pmatrix} 12 & 4 \\ -2 & 4 \end{pmatrix}$

5. $\begin{pmatrix} 0 & 4 \\ 1 & 7 \end{pmatrix}$

7. No

9. A yes, B no

11. $\begin{pmatrix} 4 & 1 & 4 \\ 7 & 1 & 0 \\ 6 & -5 & 1 \end{pmatrix}$

13. $\begin{pmatrix} 10 & 2 & 8 \\ 13 & 3 & 0 \\ 15 & -14 & 2 \end{pmatrix}$

15. $\begin{pmatrix} 0 & -9 & 0 \\ -1 & 1 & -3 \\ -4 & 0 & -1 \end{pmatrix}$

17. No, no, yes, yes

SECTION 7.2

1. $A^T B = (4 \quad -2)$, $B^T A = \begin{pmatrix} 4 \\ -2 \end{pmatrix}$

3. $C^2 = \begin{pmatrix} 1 & -3 & -3 \\ 12 & 4 & 3 \\ 0 & -4 & 1 \end{pmatrix}$

5. $BB^T = \begin{pmatrix} 4 & 0 & 2 \\ 0 & 9 & -3 \\ 2 & -3 & 2 \end{pmatrix}$, $B^T B = \begin{pmatrix} 5 & -1 \\ -1 & 10 \end{pmatrix}$

$$7. C^2 = \begin{pmatrix} 1 & -3 & -3 \\ 12 & 4 & 3 \\ 0 & -4 & 1 \end{pmatrix} \quad C^4 = \begin{pmatrix} -35 & -3 & -15 \\ 60 & -32 & -21 \\ -48 & -20 & -11 \end{pmatrix}$$

$$9. A^TCA = (5)$$

$$11. AB = BA = \begin{pmatrix} -5 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & -5 \end{pmatrix}$$

$$15. y = ABz, AB = \begin{pmatrix} \cos \theta & -2 \sin \theta \\ \sin \theta & 2 \cos \theta \end{pmatrix} \neq BA = \begin{pmatrix} \cos \theta & -\sin \theta \\ 2 \sin \theta & 2 \cos \theta \end{pmatrix}.$$

Geometrically this means that we obtain two different composite transformations, depending on the order of composition.

$$17. \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

SECTION 7.3

$$1. 16$$

$$3. 2$$

$$5. 18$$

$$7. -126$$

$$9. 90$$

$$11. 4a^2b^2c^2$$

$$13. 0$$

$$5. \begin{vmatrix} 0 & 13 & 11 \\ 34 & 20 & 6 \\ 4 & -1 & 3 \end{vmatrix}$$

SECTION 7.4

$$5. 1$$

$$7. 1$$

$$9. 2$$

SECTION 7.5

$$1. x = 2, y = 3$$

$$3. x = 5, y = 2$$

$$5. x = 1, y = 2, z = 3$$

$$7. x = 3, y = -2, z = 2$$

$$9. x = 2, y = 2, z = 2$$

$$11. I_1 = (R_2 + R_3)E_0/D, I_2 = R_3E_0/D, I_3 = R_2E_0/D, D = (R_1 + R_4)(R_2 + R_3) + R_2R_3$$

$$13. I_1 = \frac{R_4E_1 + R_5(E_1 + E_2)}{(R_1 + R_2 + R_3)(R_4 + R_5) + R_4R_5}, \quad \text{etc.}$$

SECTION 7.6

$$1. y = 2x, z = 3x, x \text{ arbitrary}$$

$$3. z = -4x, y = 0, x \text{ arbitrary}$$

$$5. y = -3x, x \text{ arbitrary}$$

$$7. x = 2z, y = 4z, z \text{ arbitrary}$$

$$9. x = 2w, y = 4w, z = 3w, w \text{ arbitrary}$$

$$11. ax + by + c = 0 \text{ (The line } L \text{ to be found.)}$$

$$ax_1 + by_1 + c = 0 \text{ (} P_1 \text{ on } L\text{.)}$$

$$ax_2 + by_2 + c = 0 \text{ (} P_2 \text{ on } L\text{.)}$$

This homogeneous system in the unknowns a, b, c has a nontrivial solution if, and only if, the determinant of the coefficients is zero; this yields the desired result.

SECTION 7.7

$$1. x = -1, y = 5z - 3$$

$$3. y = x - 1, z = 1$$

$$5. x = -z, y = 3z + 5$$

$$7. w = 2, y = x - 3z + 4$$

$$9. w = 3, x = y - 2z + 1$$

$$11. x = 1, y = 2, z = 3$$

$$13. w = 0, x = 2y, z = 3y + 1$$

SECTION 7.8

3. Dependent 5. Independent 7. Independent 9. Dependent
 11. Independent 13. Dependent 15. k arbitrary

SECTION 7.9

1. $x = 1, y = 2$ 3. $x = \frac{1}{2}, y = -\frac{1}{2}$
 5. $x = 0, y = 0, z = 0$ 7. $x = 1, y = 3, z = 2$
 9. $x = 2, y = 0, z = -2$ 11. $x = 2, y = 1, z = 3$
 13. $w = z, x = 2z, y = 3z + 2$ 15. $w = 2, x = 0, y = 1, z = 3$

SECTION 7.10

3. $\begin{pmatrix} -\frac{1}{18} & \frac{5}{18} \\ \frac{5}{18} & \frac{2}{18} \end{pmatrix}$ 5. $\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$
 7. $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ 9. $\begin{pmatrix} 2 & 0 & -1 \\ 5 & 1 & 0 \\ 0 & 1 & 3 \end{pmatrix}$ 11. $\begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix}$

13. $x = 6x^* + 4y^* + 3z^*$
 $y = 4x^* + 3y^* + 4z^*$
 $z = 3x^* + 2y^* + 2z^*$

15. $\begin{pmatrix} 0.04 & 0 & -0.01 \\ 0.01 & 0.04 & -0.01 \\ 0.03 & -0.01 & 0.01 \end{pmatrix}$, $\begin{pmatrix} 4 & -1 & -5 \\ 15 & 1 & -5 \\ 5 & 4 & 9 \end{pmatrix}$

SECTION 7.11

1. No 3. 6, 4, $\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ 5. 0, any vector 7. 10, -10, $\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \end{pmatrix}$
 9. 2.5, -2.5, $\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ 11. 3, 2, 5, $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$
 13. -20, -10, 5, $\begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$ 15. 3, 2.5, 2, $\begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$

19. This follows from the fact that the zeros of a polynomial with real coefficients are real or complex conjugates in pairs.

25. Proof by induction. For $m = 0$ and $m = 1$ the statement is true. Let $Ax_j = \lambda_j x_j$, where x_j is an eigenvector corresponding to λ_j . Suppose that the statement is true when $m = k$, that is, $A^k x_j = \lambda_j^k x_j$. By premultiplying by A we obtain

$$A^{k+1} x_j = A \lambda_j^k x_j = \lambda_j^k (Ax_j) = \lambda_j^{k+1} x_j,$$

and the proof is complete.

33. The characteristic equation is $\lambda^2 + (a + b + c)\lambda + ac = 0$ where $a = k_1/m_1$, $b = k_2/m_1$, $c = k_3/m_2$. The discriminant is

$$(a + b + c)^2 - 4ac = (a + b - c)^2 + 4bc > 0.$$

Thus the roots are real. Since their product equals $ac (>0)$ and their sum equals $-(a+b+c)$, the statement follows.

35. $y_1 = \sin \sqrt{2}t, y_2 = 2 \sin \sqrt{2}t$ 37. $y_1 = 2e^{5t} - e^{-5t}, y_2 = e^{5t} + 2e^{-5t}$

39. $y_1 = 2 \cosh 2t, y_2 = 4 \sinh 2t, y_3 = 2 \cosh 2t$

SECTION 7.12

3. $\begin{pmatrix} 0 & 2 \\ 2 & -5 \end{pmatrix}$

5. $\begin{pmatrix} -2 & 3 \\ 3 & 4 \end{pmatrix}$

7. $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

9. $\begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 1 \\ 2 & 1 & 1 \end{pmatrix}$

13. If A is given, $H = (A + \bar{A}^T)/2$ is Hermitian, $S = (A - \bar{A}^T)/2$ is skew-Hermitian, and $A = H + S$.

15. Not positive definite

17. -2 19. 6

21. $9a + 24 \operatorname{Im} b + 16c$

23. $|x_1|^2 + 2\operatorname{Re}[3x_2x_3 + ix_1x_2] + 2|x_3|^2$

25. $C^T = aA^T + bB^T = -a\bar{A} - b\bar{B}$ must equal $-\bar{C} = -a\bar{A} - b\bar{B}$. Thus $a = \bar{a}, b = \bar{b}$.
Ans. a and b real.

27. 0

29. $2i|x_1|^2 + 8i \operatorname{Im} x_1x_2$

SECTION 7.13

5. Rotation about the x_3 -axis in space

7. $\lambda = \pm 1, \begin{pmatrix} 1 \\ i - i\sqrt{2} \end{pmatrix}, \begin{pmatrix} 1 \\ i + i\sqrt{2} \end{pmatrix}$

9. Let A be unitary. Set $A^{-1} = B$. Then $B^T = (A^{-1})^T = (A^T)^{-1} = (\bar{A}^{-1})^{-1} = \bar{B}^{-1}$.

11. Let $A^T = \bar{A}^{-1}, B^T = \bar{B}^{-1}$ and set $C = AB$. Then $C^T = (AB)^T = B^T A^T = \bar{B}^{-1} \bar{A}^{-1} = (\bar{A}\bar{B})^{-1} = \bar{C}^{-1}$.

13. Let $Ax = \lambda x$ ($x \neq 0$), $Ay = \mu y$ ($y \neq 0$). Then $(Ax)^T = x^T A^T = x^T A = \lambda x^T$. Thus $\lambda x^T y = x^T A y = x^T \mu y = \mu x^T y$. Hence, if $\lambda \neq \mu$, $x^T y = 0$, which proves orthogonality.

19. trace $\bar{A} = \operatorname{trace}[T^{-1}(AT)] = \operatorname{trace}[(AT)T^{-1}] = \operatorname{trace}(ATT^{-1}) = \operatorname{trace} A$, because $\operatorname{trace} AB = \operatorname{trace} BA$.

21. $Ax = \lambda x$ ($x \neq 0$). Thus (a) $T^{-1}Ax = \lambda T^{-1}x = \lambda y$. Now $T^{-1}Ax = T^{-1}ATy = \bar{A}y$, and (a) becomes $\bar{A}y = \lambda y$.

27. $\lambda_1 = -k, \lambda_2 = -3k, x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, x_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, y = x_1 \cos \sqrt{k}t + x_2 \sin \sqrt{3k}t$

SECTION 7.14

1. $0, 0$, radii 1, 0

3. $5, 5$, radii 1, 1

5. $-9, -9, -9$, radii 1, 2, 1

11. We have (a) $B^T = (\bar{U}^{-1}\bar{A}\bar{U})^T = \bar{U}^T \bar{A}^T (\bar{U}^{-1})^T = U^{-1} \bar{A}^T U$. Thus

$$B^T B = U^{-1} \bar{A}^T U U^{-1} A U = U^{-1} \bar{A}^T A U = U^{-1} A \bar{A}^T U.$$

On the other hand, by (a), $B\bar{B}^T = U^{-1} A U U^{-1} \bar{A}^T U = U^{-1} A \bar{A}^T U = B^T B$.

13. $\sqrt{52}$ 15. $\sqrt{116}$ 17. $\sqrt{21}$

19. $26 \leq \lambda \leq 34, 26 \leq \lambda \leq 34, 28.66 \leq \lambda \leq 30$

SECTION 7.15

1. $m_0 = 1304, m_1 = 6412, m_2 = 31736, q = 4.9172, |\epsilon| \leq \sqrt{24.3374 - 24.1788} \approx 0.398$

3. $q = 3, |\epsilon| \leq \sqrt{3}$

7. $q = \frac{5}{4}, |\epsilon| \leq \frac{\sqrt{11}}{4} \approx 0.83. q = \frac{14}{9}, |\epsilon| \leq \frac{\sqrt{101}}{9} \approx 1.1$

9. $y = Ax = \lambda x, m_1 = x^T y = \lambda m_0, m_2 = \lambda^2 m_0, q = \lambda, \frac{m_3}{m_0} - q^2 = \lambda^2 - \lambda^2 = 0$ and $\epsilon =$

SECTION 8.1

1. $2\pi, 2\pi, \pi, \pi, 2\pi/n, 2\pi/n$

17. $\pi/2 (n=0), 0 (n \text{ even}), 1/n (n=1, 5, 9, \dots), -1/n (n=3, 7, 11, \dots)$

19. 0 21. $0 (n=0), 2\pi/n (n=1, 3, \dots), -2\pi/n (n=2, 4, \dots)$

23. $\pi^3/3 (n=0), (-1)^n 2\pi/n^3 (n=1, 2, \dots)$

25. $n[(-1)^n e^{-x} - 1]/(1 + n^2)$

SECTION 8.2

1. $\frac{1}{2} + \frac{2}{\pi} (\sin x + \frac{1}{3} \sin 3x + \frac{1}{5} \sin 5x + \dots)$

3. $\frac{2}{\pi} (\sin x - \frac{1}{2} \sin 2x + \frac{1}{3} \sin 3x + \frac{1}{5} \sin 5x - \frac{1}{6} \sin 6x + \dots)$

5. $2(\sin x - \frac{1}{3} \sin 2x + \frac{1}{5} \sin 3x - \frac{1}{7} \sin 4x + \dots)$

7. $2 \left[\left(\frac{\pi^2}{1} - \frac{6}{1^3} \right) \sin x - \left(\frac{\pi^2}{2} - \frac{6}{2^3} \right) \sin 2x + \left(\frac{\pi^2}{3} - \frac{6}{3^3} \right) \sin 3x - \dots \right]$

9. $\frac{4}{\pi} (\cos x + \frac{1}{3} \cos 3x + \frac{1}{5} \cos 5x + \dots)$

11. $\frac{4}{\pi} (\cos x - \frac{1}{3} \cos 3x + \frac{1}{5} \cos 5x - \dots)$

13. $\frac{2}{\pi} \sin x + \frac{1}{2} \sin 2x - \frac{2}{9\pi} \sin 3x - \frac{1}{4} \sin 4x + \frac{2}{25\pi} \sin 5x + \dots$

15. $a_0 = \pi^2/6, a_n = (-1)^n 2/n^3, b_n = \{2[(-1)^n - 1]/n^3\pi\} - \{\pi(-1)^n/n\}$

SECTION 8.3

3. Even: $|x^3|, x^2 \cos nx, \cosh x$. Odd: $x \cos nx, \sinh x, x|x|$

5. Odd

7. Neither even nor odd

9. Even

11. Odd

13. Neither even nor odd

15. $|f(-x)| = |-f(x)| = |f(x)|, f^2(-x) = (-1)^2 f^2(x) = f^2(x)$

17. $f(-x)g(-x) = [-f(x)][-g(x)] = f(x)g(x)$

21. $\cosh x + \sinh x$

23. $x^2(1-x^2)^{-1} + x(1-x^2)^{-1}$

25. Cf. Prob. 5 in Sec. 8.2.

27. Cf. Prob. 7 in Sec. 8.2.

29. $\frac{4}{\pi} (\sin x - \frac{1}{3} \sin 3x + \frac{1}{5} \sin 5x - \dots)$

31. $-\pi - \frac{4}{\pi} (\cos x + \frac{1}{3} \cos 3x + \dots) + 2(\sin x + \frac{1}{5} \sin 5x + \dots)$

SECTION 8.4

$$\frac{1}{2} + \frac{2}{\pi} \left(\cos \frac{\pi t}{2} - \frac{1}{3} \cos \frac{3\pi t}{2} + \frac{1}{5} \cos \frac{5\pi t}{2} - \dots \right)$$

7. $\frac{2}{\pi} (\sin \pi t - \frac{1}{2} \sin 2\pi t + \frac{1}{3} \sin 3\pi t - + \cdots)$

9. $-\frac{4}{\pi^3} (\cos \pi t + \frac{1}{2} \cos 3\pi t + \frac{1}{8} \cos 5\pi t + \cdots) + \frac{2}{\pi} (\sin \pi t + \frac{1}{2} \sin 3\pi t + \cdots)$

11. $\frac{1}{\pi} (\sin 4t - \frac{1}{6} \sin 12t + \frac{1}{48} \sin 20t - \frac{1}{48} \sin 28t + - \cdots)$

13. $\frac{1}{3} - \frac{4}{\pi^3} (\cos \pi t - \frac{1}{2} \cos 2\pi t + \frac{1}{6} \cos 3\pi t - \frac{1}{16} \cos 4\pi t + - \cdots)$

15. $a_0 = \frac{1}{4}(3 - e^{-2})$, $a_n = \frac{2k_n}{l_n}$, $b_n = \frac{\pi n k_n}{l_n}$ ($n = 2, 4, \dots$),

$b_n = \frac{\pi n k_n}{l_n} - \frac{2}{n\pi}$ ($n = 1, 3, \dots$), where $k_n = 1 - (-1)^n e^{-2}$, $l_n = 4 + n^2\pi^2$

SECTION 8.5

1. $f(x) = 1$

3. $\frac{1}{2} + \frac{2}{\pi} \left(\cos \frac{\pi x}{l} - \frac{1}{3} \cos \frac{3\pi x}{l} + \frac{1}{5} \cos \frac{5\pi x}{l} - + \cdots \right)$

5. $\frac{1}{2} + \frac{4}{\pi^2} \left(\cos \frac{\pi x}{l} + \frac{1}{9} \cos \frac{3\pi x}{l} + \frac{1}{25} \cos \frac{5\pi x}{l} + \cdots \right)$

7. $\frac{l^3}{4} + \frac{6l^3}{\pi^2} \left[\left(\frac{4}{\pi^2} - 1 \right) \cos \frac{\pi x}{l} + \frac{1}{2^2} \cos \frac{2\pi x}{l} + \left(\frac{4}{3^2\pi^2} - \frac{1}{3^2} \right) \cos \frac{3\pi x}{l} + \cdots \right]$

9. $\frac{2}{\pi} - \frac{4}{\pi} \left(\frac{1}{1 \cdot 3} \cos \frac{2\pi x}{l} + \frac{1}{3 \cdot 5} \cos \frac{4\pi x}{l} + \frac{1}{5 \cdot 7} \cos \frac{6\pi x}{l} + \cdots \right)$

11. $\frac{4}{\pi} \left(\sin \frac{\pi x}{l} + \frac{1}{3} \sin \frac{3\pi x}{l} + \frac{1}{5} \sin \frac{5\pi x}{l} + \cdots \right)$

13. $\frac{2}{\pi} \left(\sin \frac{\pi x}{l} + \frac{2}{2} \sin \frac{2\pi x}{l} + \frac{1}{3} \sin \frac{3\pi x}{l} + \frac{1}{5} \sin \frac{5\pi x}{l} + \frac{2}{6} \sin \frac{6\pi x}{l} + \cdots \right)$

15. $\frac{2l^3}{\pi} \left[\left(1 - \frac{4}{\pi^2} \right) \sin \frac{\pi x}{l} - \frac{1}{2} \sin \frac{2\pi x}{l} + \left(\frac{1}{3} - \frac{4}{3^2\pi^2} \right) \sin \frac{3\pi x}{l} - \frac{1}{4} \sin \frac{4\pi x}{l} + \cdots \right]$

SECTION 8.6

15. No

SECTION 8.7

1. $y = C_1 \cos \omega t + C_2 \sin \omega t + \frac{1}{\omega^2 - 1} \sin t$, the numerical values of the amplitude $A(\omega)$ of the last term being

ω	0.5	0.7	0.9	1.1	1.5	2.0	10.0
$A(\omega)$	-1.33	-0.20	-5.3	4.8	0.8	0.33	0.01

3. $y = C_1 \cos \omega t + C_2 \sin \omega t + B_1 \sin t + B_3 \sin 3t + B_5 \sin 5t$ where

ω	0.5	0.9	1.1	2.0	2.9	3.1	4.0	4.9	5.1	6.0	8.0
$B_1 = 1/(\omega^2 - 1)$	-1.33	-5.3	4.8	0.33	0.13	0.12	0.07	0.04	0.04	0.01	0.02
$B_2 = 1/(3\omega^2 - 9)$	-0.913	-0.014	-0.014	-0.01	-0.19	0.18	0.02	0.01	0.01	0.004	0.003
$B_3 = 1/25(\omega^2 - 25)$	-0.002	-0.002	-0.002	-0.002	-0.003	-0.004	-0.04	0.04	0.04	0.004	0.001

5. $y = C_1 \cos \omega t + C_2 \sin \omega t + \frac{\pi^2}{12\omega^2} - \frac{1}{\omega^2 - 1} \cos t + \frac{1}{4(\omega^2 - 4)} \cos 2t + \dots$

7. $y = C_1 \cos \omega t + C_2 \sin \omega t$

$$+ \frac{1}{2\omega^2} - \frac{1}{1 \cdot 3(\omega^2 - 4)} \cos 2t - \frac{1}{3 \cdot 5(\omega^2 - 16)} \cos 4t - \dots$$

9. $y = -\frac{K}{c} \cos t \quad 11. \quad y = \frac{1 - n^2}{D} a_n \cos nt + \frac{nc}{D} a_n \sin nt, D = (1 - n^2)^2 + n^2 c^2$

13. $y = \sum_{n=1}^{\infty} \left[\frac{(-1)^n c}{n^2 D_n} \cos nt - \frac{(-1)^n (1 - n^2)}{n^3 D_n} \sin nt \right] \text{ where } D_n = (1 - n^2)^2 + n^2 c^2$

15. $I = \sum_{n=1}^{\infty} (A_n \cos nt + B_n \sin nt), A_n = \frac{80(10 - n^2)}{\pi n^2 D_n}, B_n = \frac{800}{\pi n D_n} \text{ (n odd)},$

$A_n = 0, B_n = 0 \text{ (n even)}, D_n = (10 - n^2)^2 + 100 n^2$

SECTION 8.8

3. $F = \frac{4}{\pi} \left[\sin x + \frac{1}{3} \sin 3x + \dots + \frac{1}{N} \sin Nx \right] \text{ (N odd), cf. Ex. 1, Sec. 8.2}$

7. $F = \frac{\pi^2}{3} - 4 \left(\cos x - \frac{1}{4} \cos 2x + \frac{1}{9} \cos 3x - \dots + \frac{(-1)^{N+1}}{N^2} \cos Nx \right),$

$E^* = \frac{2\pi^5}{5} - \pi \left(\frac{2\pi^4}{9} + 16 + 1 + \frac{16}{81} + \frac{1}{16} + \dots \right)$

SECTION 8.10

9. $\frac{2}{\pi} \int_0^\infty \left[\frac{a \sin aw}{w} + \frac{\cos aw - 1}{w^2} \right] \cos xw dw$

11. $\frac{6}{\pi} \int_0^\infty \frac{2 + w^2}{4 + 5w^2 + w^4} \cos xw dw$

13. $\frac{2}{\pi} \int_0^\infty \left[\left(a^2 - \frac{2}{w^2} \right) \sin aw + \frac{2a}{w} \cos aw \right] \frac{\cos xw}{w} dw$

15. $f(ax) = \frac{1}{\pi} \int_0^\infty A(w) \cos aw dw = \frac{1}{\pi} \int_0^\infty A\left(\frac{p}{a}\right) \cos xp \frac{dp}{a} \text{ where } wa = p$

SECTION 8.11

3. $1/\sqrt{2\pi}, (\cos x)/\sqrt{\pi}, (\cos 2x)/\sqrt{\pi}, \dots$

5. $1/\sqrt{\pi}, (\sqrt{2} \cos 2x)/\sqrt{\pi}, (\sqrt{2} \cos 4x)/\sqrt{\pi}, \dots$

7. $1/\sqrt{2}, \cos \pi x, \sin \pi x, \dots$

9. $1/\sqrt{\pi}, (\sqrt{2} \cos 2x)/\sqrt{\pi}, (\sqrt{2} \sin 2x)/\sqrt{\pi}, \dots$

11. $1/\sqrt{2}, x\sqrt{3}/2, (3x^2 - 1)\sqrt{5/8}, (5x^3 - 3x)\sqrt{7/8}$

13. Set $a = ct + k.$ 15. $g_1 = 1/\sqrt{2}, g_2 = \sqrt{3/2}x, g_3 = \sqrt{5/8}(3x^2 - 1)$

19. If there is a function $h(x)$ which is orthogonal to every function of our closed set, then its Fourier constants with respect to that set are all zero, and by substituting h for f in Parseval's formula, $\|h\| = 0$. Hence, the function h does not satisfy our general assumption, and this proves completeness of our set.

SECTION 8.12

1. $\sin \frac{2n+1}{2l} \pi x, n = 0, 1, \dots$

3. $y'' + \lambda y = 0, y'(0) = 0, y'(\pi) = 0$

7. This follows from $P_0 = 1$ and the orthogonality.

9. $f = \frac{1}{4} P_0 + \frac{1}{2} P_1 + \frac{5}{16} P_2 + \dots$

11. $x^3 = \frac{3}{5} P_1 + \frac{2}{5} P_3$

17. Note that $tx - t^2/2 = x^2/2 - (x-t)^2/2$.

$$27. \int_0^\infty e^{-x} x^k L_n(x) dx = \frac{1}{n!} \int_0^\infty x^k \frac{d^n}{dx^n} (x^n e^{-x}) dx = - \frac{k}{n!} \int_0^\infty x^{n-1} \frac{d^{n-1}}{dx^{n-1}} (x^n e^{-x}) dx \\ = \dots = (-1)^k \frac{k!}{n!} \int_0^\infty \frac{d^{n-k}}{dx^{n-k}} (x^n e^{-x}) dx = 0 \text{ when } n > k.$$

SECTION 8.13

3. By (1), Sec. 3.6, with $v = 1$,

$$c_m = \frac{2}{R^2 J_1^2(\alpha_{m0})} \int_0^R x J_0\left(\frac{\alpha_{m0}}{R} x\right) dx = \frac{2}{\alpha_{m0}^2 J_1^2(\alpha_{m0})} \int_0^{\alpha_{m0}} w J_0(w) dw = \frac{2}{\alpha_{m0} J_1(\alpha_{m0})}, \\ f = 2 \left(\frac{J_0(\lambda_{10}x)}{\alpha_{10} J_1(\alpha_{10})} + \frac{J_0(\lambda_{20}x)}{\alpha_{20} J_1(\alpha_{20})} + \dots \right)$$

5. $c_m = \frac{2akJ_1(\alpha_{m0}a/R)}{\alpha_{m0} R J_1^2(\alpha_{m0})}$

7. $c_m = \frac{4J_2(\alpha_{m0})}{\alpha_{m0}^2 J_1^2(\alpha_{m0})}$

9. $c_m = \frac{2R^2}{\alpha_{m0} J_1(\alpha_{m0})} \left[1 - \frac{2J_2(\alpha_{m0})}{\alpha_{m0} J_1(\alpha_{m0})} \right]$

SECTION 9.1

21. $u = 3 \ln(x^2 + y^2)/\ln 4$

23. By integration, $u = f(y)$

25. $u_x = f(y), u = xf(y) + g(y)$

27. $u = axy + bx + cy + k$

29. $u = ax + by + c$

31. $p_y - p = 0, p = A(x)e^y, u = \int p dx = f(x)e^y + g(y) \text{ where } f(x) = \int A dx$

33. $u = B(x)e^{-y} - \frac{1}{2}x^2 - xy + C(y)$

SECTION 9.3

3. $u = 0.01 \cos t \sin x$

5. $u = k(\cos t \sin x + \cos 3t \sin 3x)$

7. $u = \frac{4}{5\pi} \left(\frac{1}{4} \cos 2t \sin 2x - \frac{1}{36} \cos 6t \sin 6x + \frac{1}{100} \cos 10t \sin 10x - \dots \right)$

9. $u = 0.12 \left(\cos t \sin x - \frac{1}{2^3} \cos 2t \sin 2x + \frac{1}{3^3} \cos 3t \sin 3x - \dots \right)$

11. $27,960/\pi^6 \approx 0.9986$

13. $u = e^{k(x+y)}$

15. $u = x^k y^k$

17. $u = ke^{x^2+y^2+c(x-y)}$

19. $u = (c_1 e^{kx} + c_2 e^{-kx})(A \cos ky + B \sin ky)$, etc.

23. This follows immediately from the Euler formulas; note that t merely plays the role of a parameter, and the series is the Fourier sine series of the constant function 1, multiplied by $A \sin \omega t$.

25. Substitute $G_n(t)$ [Prob. 24] into $u(x, t)$ in Prob. 21 and use the given initial conditions.

$$\text{Then } G_n(0) = B_n = \frac{2}{l} \int_0^l f(x) \sin \frac{n\pi x}{l} dx, \quad G_n'(0) = \lambda_n B_n + \frac{2A\omega(1 - \cos n\pi)}{n\pi(\lambda_n^2 - \omega^2)} = 0$$

$$29. u = \frac{8l^2}{\pi^3} \left(\cos c \left(\frac{\pi}{l} \right)^2 t \sin \frac{\pi x}{l} + \frac{1}{3^3} \cos c \left(\frac{3\pi}{l} \right)^2 t \sin \frac{3\pi x}{l} + \dots \right)$$

$$31. u(0, t) = 0, u(l, t) = 0, u_x(0, t) = 0, u_x(l, t) = 0$$

$$33. \beta l \approx (2n + 1) \frac{\pi}{2}, n = 1, 2, \dots \text{ (more exactly 4.730, 7.853, 10.996, \dots)}$$

$$35. \beta l \approx (2n + 1) \frac{\pi}{2}, n = 0, 1, \dots \text{ (more exactly 1.875, 4.694, 7.855, \dots)}$$

SECTION 9.4

9. $u = f_1(x) + f_2(x + y)$

11. $u = xf_1(x + y) + f_2(x + y)$

13. Probs. 8, 9, 12 hyperbolic, Probs. 10, 11 parabolic

15. $y'^2 - 2y' + 1 = (y' - 1)^2 = 0, y = x + c, \Psi(x, y) = x - y, v = x, z = x - y$

SECTION 9.5

1. The solutions of the wave equation are periodic in t .

5. Since the temperatures at the ends are kept constant, the temperature will approach a steady-state (time-independent) distribution $u_1(x)$ as $t \rightarrow \infty$, and $u_I = U_1 + (U_2 - U_1)x/l$, the solution of (1) with $\partial u / \partial t = 0$, satisfying the boundary conditions.

7. $u = \sin 0.1\pi x e^{-1.752\pi^2 t/100}$

$$9. u = \frac{40}{\pi^2} \left(\sin 0.1\pi x e^{-0.01752\pi^2 t} - \frac{1}{9} \sin 0.3\pi x e^{-0.01752(3\pi)^2 t} + \dots \right)$$

$$11. u = \frac{800}{\pi^3} \left(\sin 0.1\pi x e^{-0.01752\pi^2 t} + \frac{1}{3^3} \sin 0.3\pi x e^{-0.01752(3\pi)^2 t} + \dots \right)$$

$$15. u = \frac{\pi^2}{3} - 4 \left(\cos x e^{-t} - \frac{1}{4} \cos 2x e^{-4t} + \frac{1}{9} \cos 3x e^{-9t} + \dots \right)$$

$$17. u = \frac{\pi}{8} + \left(1 - \frac{2}{\pi} \right) \cos x e^{-t} - \frac{1}{\pi} \cos 2x e^{-4t} - \left(\frac{1}{3} + \frac{2}{9\pi} \right) \cos 3x e^{-9t} + \dots$$

SECTION 9.6

$$7. \frac{1}{\sqrt{\pi}} \int_{(a-s)/r}^{(b-s)/r} e^{-w^2} dw - \frac{1}{\sqrt{\pi}} \int_{(a+s)/r}^{(b+s)/r} e^{-w^2} dw$$

SECTION 9.8

1. c increases and so does the frequency.

5. $c\pi\sqrt{260}$ (corresponding eigenfunctions $F_{4,16}, F_{16,16}$, etc.)

$$9. B_{mn} = \frac{-8}{mlm\pi^2} ((-1)^m k_n + (-1)^n l_m) \text{ where } k_n = \begin{cases} 0 & (n \text{ even}) \\ a & (n \text{ odd}) \end{cases} \text{ and } l_m = \begin{cases} 0 & (m \text{ even}) \\ b & (m \text{ odd}) \end{cases}$$

11. $B_{mn} = 0$ (m or n even), $B_{mn} = \frac{64a^4b^2}{\pi^6 m^3 n^3}$ (m, n both odd)

15. $u = \frac{0.64}{\pi^6} \sum_{m=1}^{\infty} \sum_{\substack{n=1 \\ m, n \text{ odd}}}^{\infty} \frac{1}{m^3 n^3} \cos(\pi t \sqrt{m^2 + n^2}) \sin m\pi x \sin n\pi y$

17. $u = k \cos \pi \sqrt{5}t \sin \pi x \sin 2\pi y$

SECTION 9.9

5. $a^2 u_{x^* x^*} + c^2 u_{y^* y^*}$

SECTION 9.10

11. $u_2: r = x_1/x_2 = 0.43565, u_3: r = x_1/x_3 = 0.27789, r = x_2/x_3 = 0.63788$

19. $u = 4k \sum_{m=1}^{\infty} \frac{J_2(\alpha_m)}{\alpha_m^2 J_1^2(\alpha_m)} \cos x_m t J_0(x_m r)$

SECTION 9.11

1. $a^2 u_{x^* x^*} + b^2 u_{y^* y^*} + c^2 u_{z^* z^*}$

19. $\frac{u_1 - u_0}{\ln \frac{r_1}{r_0}} \ln r + \frac{u_0 \ln r_1 - u_1 \ln r_0}{\ln \frac{r_1}{r_0}}$

SECTION 9.12

5. $u = 1$

7. $u = \frac{2}{3}r^2 P_2(\cos \phi) + \frac{1}{3} = r^2(\cos^2 \phi - \frac{1}{3}) + \frac{1}{3}$

9. $\cos 2\phi = 2 \cos^2 \phi - 1, 2x^2 - 1 = \frac{4}{3}P_2(x) - \frac{1}{3}, u = \frac{4}{3}r^2 P_2(\cos \phi) - \frac{1}{3}$

11. $u = 4r^3 P_3(\cos \phi) - 2r^2 P_2(\cos \phi) + P_1(\cos \phi) - 2$

SECTION 10.1

1. $\frac{2}{5}$

3. $x^3 - 3xy^2$

5. $\frac{4}{13}$

7. $-2xy/(x^2 + y^2)^2$

9. 1

11. 1

13. $\pi/4$

15. $-\pi/2$

21. Consider $c = a + ib = z_1/(z_1 + z_2)$, assuming $z_1 + z_2 \neq 0$. From (1), $|a| \leq |c|$, $|a - 1| \leq |c - 1|$. Thus $|a| + |a - 1| \leq |c| + |c - 1|$. Clearly $|a| + |a - 1| \geq 1$, and the last inequality becomes

$$1 \leq |c| + |c - 1| \quad \left| \frac{z_1}{z_1 + z_2} \right| + \left| \frac{z_2}{z_1 + z_2} \right|$$

Multiplication by $|z_1 + z_2|$ yields (2).

23. By (2), $|z_1| = |(z_1 + z_2) + (-z_2)| \leq |z_1 + z_2| + |z_2|$. Subtract $|z_2|$ on both sides.

25. Square both sides of both inequalities.

27. $y \geq -1$

29. The regions to the right of the right branch and to the left of the left branch of the hyperbola $x^2 - y^2 = 1$

31. Vertical strip bounded by $x = -\pi$ and $x = \pi$

33. $(x - \frac{5}{3})^2 + y^2 = \frac{16}{9}$

35. The y -axis

SECTION 10.2

1. $11 + 13i, -27 + 3i, 41 - 23i$

3. $2 - i, (4 - 3i)/5, (8 - i)/13$

5. $(1 - x)/[(1 - x)^2 + y^2], y/[(1 - x)^2 + y^2]$

7. $\operatorname{Re} w > 0$

$\operatorname{Re} |w| > 9$

11. $\operatorname{Re} w \geq 0$

15. No, because $f(z) \rightarrow 0$ [$= f(0)$] as $z \rightarrow 0$ along the y -axis, but $f(z) \rightarrow 1$ [$\neq f(0)$] as $z \rightarrow 0$ along the positive x -axis.

17. Yes, because for $z \neq 0$, $|f| = z^2/\sqrt{x^2 + y^2} \leq |z| \leq |z|$ and, therefore, $|f| \rightarrow 0$ [$= f(0)$] as $|z| \rightarrow 0$.

19. By definition of continuity, for any $\epsilon > 0$ there is a $\delta > 0$ so that $|f(z) - f(a)| < \epsilon$ when $|z - a| < \delta$. Since $\lim_{n \rightarrow \infty} z_n = a$, for all sufficiently large n the points z_n lie in the disk $|z - a| < \delta$, i.e., $|f(z_n) - f(a)| < \epsilon$ for all these n .

23. $4z(z^2 + 1)$

25. $(2 - 2z^2)/(z^2 + 1)^2$

27. $(6 - 8i)/25$

29. $-1 - 3i$

SECTION 10.3

1. Yes

3. No

5. Yes, except at $z = 1$

7. Yes, except at $z = 0$

9. No

11. a

13. $-1/z^2$

15. Let $f = u + iv$ and $|f| = c$. Then $u^2 + v^2 = c^2$. By differentiation, $uu_x + vv_x = 0$, $uu_y + vv_y = 0$. By (5), $uu_x - vu_y = 0$, $uu_y + vu_x = 0$. Hence

$$(uu_x - vu_y)^2 + (uu_y + vu_x)^2 = (u^2 + v^2)(u_x^2 + u_y^2) = c^2(u_x^2 + u_y^2) = 0.$$

If $c = 0$, then $u = v = 0$. If $c \neq 0$, then $u_x^2 + u_y^2 = 0$, i.e., $u_x = u_y = 0$ and, by (5), $v_x = 0$, $v_y = 0$. Hence, $u = \text{const}$, $v = \text{const}$.

17. Let $f = u + iv$. By (3), $f' = u_x + iv_x = 0$. Hence, $u_x = 0$, $v_x = 0$. By (4) or (5), $u_y = 0$, $v_y = 0$. Hence, $u = \text{const}$, $v = \text{const}$, $f = u + iv = \text{const}$.

23. $z^2/2$

25. $e^x(\cos y + i \sin y)$

27. $\ln r + i\theta$

29. $\cos x \cosh y - i \sin x \sinh y$

SECTION 10.4

1. $\pm(1 + i)/\sqrt{2}$

3. $2i, -2i$

5. $-1, (1 \pm i\sqrt{3})/2$

7. $i, (\pm\sqrt{3} - i)/2$

9. $\pm(1 + i)/\sqrt{2}, \pm(-1 + i)/\sqrt{2}$

11. $\pm(\sqrt{3} + i)/2, \pm i, \pm(\sqrt{3} - i)/2$

13. $3, 3(-1 \pm i\sqrt{3})/2$

15. $\pm 1, \pm(1 + i)\sqrt{2}$

17. $\pm(1 + i)\sqrt{2}$

19. $\pm(2 - i)$

21. $\pm(2 + 3i)$

23. $i, -1 - i$

25. $3 + 2i, 2 - i$

SECTION 10.5

3. From (6), $e^z e^{-z} = e^0 = 1$, which implies $e^z \neq 0$.

5. $(1 + i)/\sqrt{2}$

7. $(-1 + i)/\sqrt{2}$

9. $e(\cos 1 + i \sin 1)$

11. $e^{2x} \cos 2y, e^{2x} \sin 2y$

13. $e^{x^2-y^2} \cos 2xy, e^{x^2-y^2} \sin 2xy$

15. $e^{x^2+z-y^2} \cos(2xy + y), e^{x^2+z-y^2} \sin(2xy + y)$

17. $5e^{-z} \arctan(\frac{y}{5})$

19. $\pm 2n\pi i$ ($n = 0, 1, \dots$)

21. $\ln 2 \pm (2n + 1)\pi i$ ($n = 0, 1, \dots$)

23. $-e^{-x} \sin y$

25. $-e^z \sin y$

SECTION 10.6

17. $\sqrt{\sin^2 x + \sinh^2 y}$

19. $\sqrt{\frac{\sin^2 x + \sinh^2 y}{\cosh^2 x + \sinh^2 y}}$

21. $-\frac{\cos x \sinh y}{\sin^2 x + \sinh^2 y}$

23. $\frac{\sin x \cos x}{\sin^2 x + \sinh^2 y}$

25. (a) $\sin x \sinh y = 0$, $y = 0$ or $x = n\pi$, $n = 0, \pm 1, \pm 2, \dots$ (parallel to the y -axis).

27. $3.992 + 6.218i$

29. $2.333 - 0.274i$

31. $1.175i$

33. $1.960 + 3.166i$ 35. $\pm 2n\pi \pm 2.29i$ 37. $0, \pm \pi i, \pm \frac{2}{3}\pi i, \dots$
 39. $\pm(\pi/3)i \pm 2n\pi i, n = 0, 1, \dots$

SECTION 10.7

5. $\pm 2n\pi i, n = 0, 1, \dots$ 7. $i(\frac{1}{8}\pi \pm 2n\pi), n = 0, 1, \dots$
 9. $1 + (\frac{1}{2}\pi \pm 2n\pi)i, n = 0, 1, \dots$ 11. $(1 \pm 2n\pi)i, n = 0, 1, \dots$
 13. $-i$ 15. $-e$ 17. $0.693 + 0.785i$
 19. $\ln 6 + \frac{1}{3}\pi i \approx 1.792 + 1.047i$ 21. $e^{-\pi/4} [\cos(\ln \sqrt{2}) + i \sin(\ln \sqrt{2})]$
 23. $\sqrt{2} e^{\pi/4} [\cos(\ln \sqrt{2} - \frac{1}{4}\pi) + i \sin(\ln \sqrt{2} - \frac{1}{4}\pi)]$
 27. $\sinh w = \frac{1}{2}(e^w - e^{-w}) = z, (e^w)^2 - 2ze^w - 1 = 0, e^w = z + \sqrt{z^2 + 1}$
 29. $\tanh w = \frac{e^w - e^{-w}}{e^w + e^{-w}} = z, e^w(1 - z) = e^{-w}(1 + z), e^{2w} = \frac{1+z}{1-z}, \text{etc.}$

SECTION 11.1

1. The positive and the negative v -axis, respectively
 5. The parabola $v = \frac{1}{2}(1 - u^2)$ 7. $|w| > 4$ 9. $|\arg w| < \pi/2$
 11. The region to the right of the parabola $v^2 = 4(1 + u)$
 13. $4 < |w| < 9$ and $|\arg w| < \pi/2$
 15. $0 \leq \arg w \leq \pi/4, \pi \leq \arg w \leq 5\pi/4, -\pi/2 \leq \arg w \leq -\pi/4, 0 \leq \arg w \leq \pi/2,$
 $\pi/2 \leq \arg w \leq \pi, \pi \leq \arg w \leq 3\pi/2, -\pi/2 \leq \arg w \leq 0, 0 \leq \arg w \leq 3\pi/4$
 17. $-1 < u < 1, v > 1$
 21. $z = c$ maps onto the circle $(u - 1/2c)^2 + v^2 = 1/4c^2$, etc.; the level curves are circles.
 25. z^3 maps the given region onto the upper half plane. Multiplication by i rotates this
 half plane through the angle $\pi/2$, and addition of 1 carries the origin into the point 1.
 Thus, $w = iz^3 + 1$.

SECTION 11.2

3. No, the size is preserved, but the sense is reversed.
 5. $z = (2n + 1)\pi/2, n = 0, \pm 1, \pm 2, \dots$ 7. $z = 0$
 9. $\pm 1, \pm i$ 11. $z = 0, z = \pm 1/\sqrt{2}$ 17. $z(t) = 2 \cos t + 2i \sin t$
 19. $z(t) = t + 4it^2$

SECTION 11.3

3. $z = 2(1 - i)$ 5. $z = i + \frac{1}{2} \pm \sqrt{\frac{1}{4} + i}$ 7. $0, -i$ 9. ± 2
 11. $w = 1/z, w = (2z + 1)/(z + 2)$, etc. 13. $w = (2z - 1)/z$, etc.
 15. $w = (az + b)/(a - bz)$ 17. $u = -\frac{1}{2}$ 19. $v = -\frac{1}{4}$
 21. $|w - \frac{1}{2}| = \frac{1}{2}$
 23. The image is bounded by the circles $|w - \frac{1}{2}| = \frac{1}{2}$ and $|w - \frac{1}{4}| = \frac{1}{4}$.
 25. $w_1 = iz, w_2 = w_1 + 4, w_3 = 1/w_2, w_4 = 5iw_3, w = w_4 - i$

SECTION 11.4

3. $z = (-4w + 1)/(2w - 1)$ 7. $w = 1/z$
 9. $w = 1/z$ 11. $w = 1/(z + 1)$ 13. $w = (z - i)/(z + i)$
 15. $w = (z + 1)/(z - i)$ 17. $w = (z + 3)/(2z - 1)$ 19. $w = (2z + 3)/(3z - 2)$

SECTION 11.5

1. $|w| \leq 8, 0 \leq \arg w \leq 3\pi/4$ 3. $e^{-1} < |w| < e, |\arg w| < \pi/2$
 5. $1 < |w| < e, 0 < \arg w < 1$ 7. $e^{-8} \leq |w| \leq e^8, -\pi \leq \arg w \leq -\pi/2$

9. $t = z^2$ maps R onto the strip $0 < \operatorname{Im} t < \pi$, and $w = e^t$ maps this strip onto the upper half plane. Ans. $w = e^{z^2}$

11. The region in the upper half-plane bounded by the ellipses

$$\cosh^2 1 - \sinh^2 1 = 1 \quad \text{and} \quad \frac{-}{\cosh^2 2} + \frac{\cdot}{\sinh^2 2} = 1$$

13. Elliptical annulus bounded by the ellipses in Prob. 11 and cut along the positive imaginary axis

15. $w' = \cos z = 0$ at $z = \pm(2n + 1)\frac{\pi}{2}$, $n = 0, 1, \dots$

SECTION 11.6

1. w moves once around the unit circle $|w| = 1$.
3. 4 and 5 sheets, respectively, branch point at $z = 0$
7. 2 sheets, branch point at $z = 0$
9. 3 sheets, branch point of the third order at $z = i$

SECTION 12.1

1. $z = (1 + 2i)t$, $0 \leq t \leq 1$
3. $z = 1 + i + (2 - 5i)t$, $0 \leq t \leq 1$
5. $z = -2 + i + it$, $0 \leq t \leq 3$
7. $z = 5e^{it}$, $0 \leq t \leq 2\pi$
9. $z = t + it^2$, $0 \leq t \leq 3$
11. $z = 3 \cos t + i \sin t$, $0 \leq t \leq 2\pi$
13. $z = t + i/t$, $1 \leq t \leq 4$
15. Straight line segment from 1 to $3 - i$
17. Upper half of the circle $x^2 + (y - 3)^2 = 9$
19. $y = 3x^2$ from $(-1, 3)$ to $(2, 12)$
21. $-8i$
23. $-115 - 46i$
25. $-702 - 486i$
27. 0
29. (a) i , (b) $2i$, (c) $2i$

SECTION 12.2

5. $2b(1 + i)$
7. $\frac{4}{3}$
9. $-1 + ie$
11. $i(\sinh 2\pi - \sinh \pi)$
13. $1 - \cosh 1$
15. $i \sin 1$
17. $8\sqrt{8}$
19. $\sqrt{680}$

SECTION 12.3

3. If $z = 0$ lies exterior to C .
5. Does not apply
7. Does not apply
9. $-2(1 - i)/3$
11. $2\pi i, 0$
13. $2\pi i, \pi i$
15. 0
17. $-2\pi i, 0$
19. 0

SECTION 12.4

1. $2 + 2i$
3. $(10 - 2i)/3$
5. $-2e$
7. $(e^{-1} - 1)/2$
9. $i \sinh \pi$
11. $-\frac{1}{2} \sin \pi^2$
13. $(\pi - \frac{1}{2} \sinh 2\pi)i$
15. $i \sin 3$
17. $\cos 2 - 1$
19. 0

SECTION 12.5

1. π
3. 0
5. $\pi i/2$
7. $\pi/2$
9. $2\pi i$
11. $\pi i/2$
13. 0
15. $2\pi i$
17. 0
19. $2\pi i$

SECTION 12.6

1. $\pi i/2$
3. $3\pi/8$
5. $\pi i/8$
7. πi
9. $\pi(\pi + 4i)(1 - i)/32\sqrt{2}$
11. $2\pi i$
13. $(-1)^n 2\pi i/(2n)!$
15. $2\pi i$

SECTION 13.1

1. $\frac{1}{2}, \frac{2}{3}, \frac{3}{5}, \frac{4}{7}, \dots$
3. $i, -\frac{1}{2}, -i/27, \frac{1}{3}\sqrt{3}, \dots$
5. $i, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \dots$
7. $1, 2, 3, 2, 1, 2, 3, \dots$

9. $\frac{1}{3}, \frac{1}{4}, \frac{4}{3}, \frac{3}{16}, \frac{54}{5}, \frac{127}{1024}, \dots$
 13. Bounded, convergent, limit i
 17. Bounded, convergent, limit πi
 21. Unbounded, divergent, not monotone, limit point 0
 23. Bounded, convergent, limit 1, not monotone
 25. Unbounded, divergent, monotone
 27. Bounded when $|a| \leq 1$, convergent when $-1 < a < 1$ (limit 0) and when $a = 1$ (limit 1), monotone when $a \geq 0$. Limit points 1 and -1 when $a = -1$
 29. Bounded, convergent, limit 0, monotone when $0 \leq c \leq \frac{1}{4}$

SECTION 13.3

- | | | |
|----------------|---------------|----------------------------------|
| 1. Convergent | 3. Divergent | 5. Convergent [compare with (5)] |
| 7. Convergent | 9. Convergent | 11. Divergent |
| 15. Convergent | 17. Divergent | 19. Convergent |
| 23. 2 | 25. 5 | 21. 4, 8, 66 |

SECTION 13.5

- | | | | |
|-------------------|--------------|-------------------|-------------------|
| 3. 1 | 5. 3 | 7. ∞ | 9. ∞ |
| 11. $\frac{1}{4}$ | 13. ∞ | 15. $\frac{1}{6}$ | 17. $\frac{1}{6}$ |
| | | | 19. 4 |

SECTION 13.6

7. 1 9. 1 11. 3

SECTION 14.2

7. $1 - \frac{(2z)^2}{2!} + \frac{(2z)^4}{4!} - + \dots, R = \infty$ 9. $1 - z + \frac{z^2}{2!} - \frac{z^3}{3!} + - \dots, R = \infty$
 11. $-1 - (z - \pi i) - \frac{1}{2!}(z - \pi i)^2 - \frac{1}{3!}(z - \pi i)^3 - \dots, R = \infty$
 13. $\frac{1}{\sqrt{2}} \left[1 + \left(z + \frac{\pi}{4} \right) - \frac{1}{2!} \left(z + \frac{\pi}{4} \right)^2 - \frac{1}{3!} \left(z + \frac{\pi}{4} \right)^3 + \dots \right], R = \infty$
 15. $-1 - (z + 1) - (z + 1)^2 - (z + 1)^3 - \dots, R = 1$
 17. $\cos^2 z = \frac{1}{2} (1 + \cos 2z) = \frac{1}{2} \left[2 - \frac{2^2}{2!} z^2 + \frac{2^4}{4!} z^4 - + \dots \right], R = \infty$
 19. $z + \frac{1}{3} z^3 + \frac{2}{15} z^5 + \dots, R = \frac{\pi}{2}$ 21. $z + z^3 + \frac{z^5}{3} + \dots, R = \infty$
 27. $z + \frac{z^3}{3} + \frac{z^5}{2! 5} + \frac{z^7}{3! 7} + \dots, R = \infty$ 29. $z - \frac{z^5}{2! 5} + \frac{z^9}{4! 9} - \frac{z^{13}}{6! 13} + - \dots, R = \infty$

SECTION 14.3

3. $1 - z^3 + z^6 - + \dots, |z| < 1$ 5. $1 - 2z^2 + 3z^4 - 4z^6 + - \dots, |z| < 1$
 7. $1 - \frac{z^4}{2!} + \frac{z^8}{4!} - \frac{z^{12}}{6!} + - \dots$ 9. $1 + z^4 + \frac{z^8}{2!} + \frac{z^{12}}{3!} + \dots$
 11. $\frac{z^8}{1! 3} - \frac{z^4}{3! 7} + \frac{z^6}{5! 11} - \frac{z^8}{7! 15} + - \dots, |z| < \infty$

15. $e[1 + z + \frac{1}{2}z^2 + \frac{1}{3}z^3 + \dots], |z| < \infty$

17. $e[1 + z + \frac{3}{2}z^2 + \frac{13}{6}z^3 + \dots], |z| < 1$

19. $\frac{1}{1-3i} + \frac{3}{(1-3i)^2}[z - (1+i)] + \frac{3^3}{(1-3i)^3}[z - (1+i)]^3 + \dots,$
 $|z - (1+i)| < \frac{1}{3}\sqrt{10}$

21. $\frac{i}{2} \left(1 - (1+i)(z+i) + \frac{3}{4}(1+i)^2(z+i)^2 - \frac{(1+i)^3}{2}(z+i)^3 + \dots \right), |z+i| < \sqrt{2}$

23. $1 + 2\left(z - \frac{\pi}{4}\right) + 2\left(z - \frac{\pi}{4}\right)^2 + \frac{8}{3}\left(z - \frac{\pi}{4}\right)^3 + \dots, \left|z - \frac{\pi}{4}\right| < \frac{\pi}{4}$

SECTION 14.4

1. $n = 7, 10, 16, 27, 65$

7. Convergence follows from Theorem 1, Sec. 13.3. Let $R_n(z)$ and R_n^* be the remainders of (1) and (5), respectively. Since (5) converges, for given $\epsilon > 0$ we can find an $N(\epsilon)$ such that $R_n^* < \epsilon$ for all $n > N(\epsilon)$. Since $|f_n(z)| < M_n$ for all z in the region R we have $|R_n(z)| < R_n^*$ and therefore $|R_n(z)| < \epsilon$ for all $n > N(\epsilon)$ and all z in the region R . This proves that the convergence of (1) in the region R is uniform.

9. $s_n = x^n; s = \lim_{n \rightarrow \infty} s_n = \begin{cases} 0 & \text{when } 0 \leq x < 1 \\ 1 & \text{when } x = 1 \end{cases}$ is discontinuous at $x = 1$ and, by Theorem 2, the statement follows.

11. This follows from Theorem 1.

13. $|\tanh z| \leq 1, |\tanh^n z| \leq 1, \sum 1/n!$ converges.

15. $\frac{1}{|z| + n^{\frac{1}{2}}} \leq \frac{1}{n^{\frac{1}{2}}}$

17. $\left| \frac{z^n}{n^{\frac{1}{2}}} \right| = \frac{|z|^n}{n^{\frac{1}{2}}} \leq \frac{1}{n^{\frac{1}{2}}} \text{ in } R$

SECTION 14.5

1. $\frac{1}{z^3} - \frac{1}{z^2} + \frac{1}{2z} - \frac{1}{6} + \frac{z}{24} - + \dots, R = \infty$

3. $\frac{1}{z^3} - 2 + \frac{2}{3}z^3 - \frac{4}{45}z^4 + - \dots, R = \infty$

5. $\frac{1}{z^3} + 1 + z^3 + z^6 + z^9 + \dots, R = 1$

7. $\frac{3}{z^3} + \frac{9}{2} + \frac{81}{40}z^3 + \dots, R = \infty$

9. Let $\sum_{-\infty}^{\infty} A_n(z-a)^n$ and $\sum_{-\infty}^{\infty} B_n(z-a)^n$ be two Laurent series of the same function $f(z)$ in the same annulus. We multiply both series by $(z-a)^{-k-1}$ and integrate along a circle with center at a in the interior of the annulus. Since the series converge uniformly, we may integrate term by term. This yields $2\pi i A_k = 2\pi i B_k$. Thus, $A_k = B_k$ for all $k = 0, \pm 1, \dots$.

11. $\sum_{n=0}^{\infty} \left(\frac{-4}{n} \right) (z-1)^n, (|z-1| < 1), \quad \sum_{n=0}^{\infty} \left(\frac{-4}{n} \right) \frac{1}{(z-1)^{n+4}}, (|z-1| > 1)$

13. $\frac{2}{z} + \sum_{n=0}^{\infty} \frac{(-1)^n - 4}{3^{n+1}} z^n, (0 < |z| < 3), \quad \frac{2}{z} + \sum_{n=0}^{\infty} \frac{3^n(4 + (-1)^n)}{z^{n+1}}, (|z| > 3)$

15. $\sin z = \sin \left[\left(z - \frac{\pi}{4} \right) + \frac{\pi}{4} \right] = \frac{1}{\sqrt{2}} \left[\sin \left(z - \frac{\pi}{4} \right) + \cos \left(z - \frac{\pi}{4} \right) \right] +$
 $\cdot \frac{1}{\sqrt{2}} \left[\frac{1}{\left(z - \frac{\pi}{4} \right)^3} + \frac{1}{\left(z - \frac{\pi}{4} \right)^2} - \frac{1}{2! \left(z - \frac{\pi}{4} \right)} - \frac{1}{3!} + \frac{z - \frac{\pi}{4}}{4!} + \dots \right], \quad \left| z - \frac{\pi}{4} \right| > 0$

SECTION 14.6

1. No

3. No

5. No

7. No

SECTION 15.1

1. ± 1 (simple) 3. $-i$ (third order) 5. $0, \pm 1, \pm 2, \dots$ (third order)7. $(2n+1)\pi i/2, n = 0, \pm 1, \pm 2, \dots$ (second order)9. $\pm 2n\pi i, n = 0, 1, \dots$ (simple) 11. $\pm 1/n\pi$ (double)13. $0, \infty$ (simple poles)15. ∞ (pole of second order)17. $-a$ (pole of third order)19. $0, \infty$ (essential singularities)21. ∞ (essential singularity)23. 0 (essential singularity)25. 0 (essential singularity), 1 (simple pole)27. $-i, i$ (poles of second order), ∞ (essential singularity)29. If $f(z)$ has a pole of m th order at $z = a$, then

$$f(z) = \frac{c_m}{(z-a)^m} + \frac{c_{m-1}}{(z-a)^{m-1}} + \dots = \frac{c_m}{(z-a)^m} \left[1 + \frac{c_{m-1}}{c_m}(z-a) + \dots \right], \quad c_m \neq 0.$$

For given $M > 0$, no matter how large, we can find a $\delta > 0$ so small that $\frac{|c_m|}{\delta^m} > 2M$ and

$\left| \left[1 + \frac{c_{m-1}}{c_m}(z-a) + \dots \right] \right| > \frac{1}{2}$ for all $|z-a| < \delta$. Then $|f| > \frac{|c_m|}{\delta^m} \frac{1}{2} > M$. Hence
 $|f| \rightarrow \infty$ as $z \rightarrow a$.

SECTION 15.2

1. 1

3. 0

5. $-\frac{1}{2}$ (at $z = -1$), $\frac{1}{2}$ (at $z = 1$)7. $\frac{1}{4}, \frac{1}{2}, -\frac{1}{4}, -\frac{1}{2}$ (at $z = 1, -1, i, -i$)9. $(-1)^n$ (at $z = \pm n\pi, n = 0, 1, \dots$)11. 1 (at $z = \pm n\pi$)13. 1 (at $z = 0$)15. 1 (at $z = 0$) 17. 1 (at $z = 0$)19. 0 (at $z = \pm n\pi$)21. $-\frac{1}{16}, \frac{1}{12}, \frac{5}{48}$ (at $z = -4, -2, 4$)23. $-1/2\pi$ (at $z = 3\pi$), $1/2\pi$ (at $z = \pi$)25. 0 (at $z = 0$) 27. $2\pi i$ 29. $\pi i/3$ 31. $2\pi i$ 33. $2\pi i$

35. 0

37. 0

SECTION 15.3

1. $2\pi i$

3. 0

5. $5\pi i$ 7. $2\pi i/3$ 9. $-\pi i$

11. 0

13. $2\pi i/9$ 15. $-16\pi i/9$ 17. $-4i$ 19. $-4i$ 21. $-2\pi i/5$

23. 0

25. $-2\pi/3$ 27. $-\pi \sin 0.5$

29. 0

SECTION 15.4

1. $\frac{2\pi}{\sqrt{3}}$ 3. $\frac{\pi}{\sqrt{k^2 - 1}}$ 5. $\frac{\pi}{2}$

7. 0

9. $\frac{2\pi}{1-k^2}$

11. $\cos 2\theta = \frac{1}{2} \left(z^2 + \frac{1}{z^2} \right)$, $\int_0^{2\pi} \frac{\cos^2 \theta}{26 - 10 \cos 2\theta} d\theta = -\frac{1}{20i} \int_C \frac{(z^2 + 1)^2}{z(z^2 - \frac{1}{z})(z^2 - 5)} dz = \frac{\pi}{20}$

15. $2\pi/3$ 17. $3\pi/8$ 19. 0 21. $\pi/12$ 23. $\pi/2$

25. The integrals are zero because the integrands are odd. 27. π/e

29. $\int_{-a}^a e^{-x^2} dx + i e^{-a^2} \int_0^b e^{y^2 - 2ay} dy + e^{b^2} \int_a^{-a} e^{-x^2 - 2bx} dx + i e^{-b^2} \int_b^0 e^{y^2 + 2ay} dy = 0.$

Let $a \rightarrow \infty$. Then the terms having the factor e^{-a^2} approach zero. Taking the real part of the third integral, we thus obtain

$$-\int_{-\infty}^{\infty} e^{-x^2} \cos 2bx dx = 2 \int_0^{\infty} e^{-x^2} \cos 2bx dx = e^{-b^2} \int_{-\infty}^{\infty} e^{-x^2} dx = e^{-b^2} \sqrt{\pi}.$$

Ans. $\frac{1}{2}e^{-b^2}\sqrt{\pi}$.

SECTION 16.1

1. $u(r) = 100 \ln r / \ln 5 = 62.13 \ln r$ 3. $u(r) = 100(\ln r + \ln 5) / \ln 10$
 5. $(x - 1/2c)^2 + y^2 = 1/4c^2$ 7. $u = c \operatorname{Re} \ln(z^2 - a^2) = c \ln |z^2 - a^2|$

SECTION 16.3

3. 0, $\frac{3}{4}$, $\frac{5}{8}$, $\frac{1}{2}$, $1 - e^{\frac{1}{2}i\theta}$
 5. For real k , this is a parallel flow in the y -direction, and $V = ik$.
 7. $V = 2x + 1 - 2iy$, $V_z = 0$ on the z -axis
 9. $V = 3(x^2 - y^2) - 6xyi$, $V_z = 0$ on the coordinate axes
 11. $V = 4x(x^2 - 3y^2) + 4iy(y^2 - 3x^2)$, $V_z = 0$ on the z -axis and on the lines $y = \pm \sqrt{3}x$
 13. $F(z) = (\ln(z + a))/2\pi$
 17. The streamlines are circles $y/(x^2 + y^2) = c$ or $x^2 + (y - k)^2 = k^2$.

SECTION 16.4

3. $|f| = e^z$ is monotone. Hence its minimum lies at that point z (or those points) of R which has the smallest value of z among all points in R ; clearly this is a point on the boundary. Similarly for the maximum.

5. This follows immediately from Theorem 3.

SECTION 16.5

3. $u = r \sin \theta$ 5. $u = r^3 \sin 3\theta$
 7. $u = 3r \sin \theta - r^3 \sin 3\theta$ 9. $u = \frac{1}{2} + \frac{2}{\pi} \left(r \sin \theta + \frac{r^3}{3} \sin 3\theta + \frac{r^5}{5} \sin 5\theta + \dots \right)$
 11. $u = \left(1 + \frac{2}{\pi} \right) r \sin \theta - \frac{r^3}{2} \sin 3\theta + \left(\frac{1}{3} - \frac{2}{9\pi} \right) r^3 \sin 3\theta - \frac{r^4}{4} \sin 4\theta + \dots$

SECTION 17.1

1. 1.167 3. -3.723 5. -0.931 7. 0.5
 9. 0.05 11. 1.077

SECTION 17.2

3. $C(x) = x - \frac{x^5}{215} + \frac{x^9}{419} - \frac{x^{13}}{6113} + \dots$

5. $\text{Si}(x) = x - \frac{x^3}{3!3} + \frac{x^5}{5!5} - \frac{x^7}{7!7} + \dots$

7. At the zeros of the integrands. 9. Set $t^2 = \tau$ in (7).

11. Consider $C(x) + iS(x)$ and $C(x) - iS(x)$.

13. Since $\text{Si}(x)$ is odd, we may consider it for $x \geq 0$. Since $\sin x$ is periodic and $1/x$ decreases in a monotone fashion, we have

$$\left| \int_0^\pi \frac{\sin t}{t} dt \right| > \left| \int_\pi^{2\pi} \frac{\sin t}{t} dt \right| > \left| \int_{2\pi}^{3\pi} \frac{\sin t}{t} dt \right| > \dots$$

Since the first integral is positive, the second is negative, the third is positive, etc., the statement follows.

15. The integrands are entire. Apply Theorem 4 in Sec. 13.6.

SECTION 17.3

3. Consider the quotient $q_n = \left| \frac{c_n/x^n}{c_{n+1}/x^{n+1}} \right| = \left| \frac{c_n x}{c_{n+1}} \right|$. For n for which $q_n > 1$ the terms

are still decreasing in absolute value until an n is reached for which $q_n < 1$ and the terms increase. Hence, the smallest n for which $q_n \leq 1$ is the subscript of the smallest term in absolute value. If some of the c_n are zero, the method must be modified in an obvious fashion.

5. The term involving x^{-3}, x^{-9}, x^{-19} , respectively

7. The term involving x, x^2, x^{n-1} , respectively 9. Use (1).

SECTION 17.4

1. Cf. (7) in Sec. 17.3.

3. y satisfies $y' = x^{-1}(1 - \alpha + x)y - 1$. Ans.

$$Q(\alpha, x) \sim x^{\alpha} e^{-x} \left[\frac{1}{x} - \frac{1-\alpha}{x^2} + \dots + (-1)^{n-1} \frac{(1-\alpha)(2-\alpha)\dots(n-1-\alpha)}{x^n} + \dots \right].$$

5. $\text{ci}(x) \sim \sin x \left(-\frac{1}{x} + \frac{2!}{x^3} - \frac{4!}{x^5} + \dots \right) + \cos x \left(\frac{1}{x^2} - \frac{3!}{x^4} + \frac{5!}{x^6} - \dots \right)$

7. $s(x) \sim \frac{1}{2} \left(\frac{1}{x} - \frac{1 \cdot 3}{4x^3} + \frac{1 \cdot 3 \cdot 5 \cdot 7}{16x^5} - \dots \right) \cos x^2$

$$+ \frac{1}{2} \left(\frac{1}{2x^3} - \frac{1 \cdot 3 \cdot 5}{8x^7} + \dots \right) \sin x^2$$

9. Cf. Prob. 3.

11. $C(x) \sim K + A(x) \sin x^2 - B(x) \cos x^2$, $S(x) \sim K - A(x) \cos x^2 - B(x) \sin x^2$ where $K = \pi^{1/2} 2^{-3/4}$ and

$$A(x) = \frac{1}{2} \left(\frac{1}{x} - \frac{1 \cdot 3}{4x^3} + \frac{1 \cdot 3 \cdot 5 \cdot 7}{16x^5} - \dots \right), B(x) = \frac{1}{2} \left(\frac{1}{2x^3} - \frac{1 \cdot 3 \cdot 5}{8x^7} + \dots \right).$$

13. Set $p = -\ln t$. Then $t = e^{-p}$, $dt = -e^{-p} dp$, and

$$\text{li}(e^{-x}) = \int_0^{-x} \frac{dt}{\ln t} = \int_{\infty}^x \frac{-e^{-p}}{-p} dp = - \int_x^{\infty} \frac{e^{-p}}{p} dp = -\text{Ei}(x)$$

SECTION 18.3

1. $\bar{x} = 83.3$, $s^2 = 13.17$

3. $\bar{x} = 3.47$, $s^2 = 2.98$

5. $\bar{x} = 99.2$, $s^2 = 234.7$, grouped $\bar{x} = 99.4$, $s^2 = 254.7$

7. 360

9. 96, 86, 109.5, 23.5

11. $\bar{x} = \frac{1}{n} \sum x_j = \frac{1}{n} \sum (x_j^* + c) = \frac{1}{n} \sum x_j^* + c = \bar{x}^* + c$,

$$\begin{aligned}s^2 &= \frac{1}{n-1} \sum [x_j - \bar{x}]^2 = \frac{1}{n-1} \sum [(x_j^* + c) - (\bar{x}^* + c)]^2 \\ &= \frac{1}{n-1} \sum [x_j^* - \bar{x}^*]^2 = s^{*2}\end{aligned}$$

SECTION 18.4

1. $A \cup B$ occurs if and only if $A \cap B$ or $A \cap \bar{B}$ or $\bar{A} \cap B$ occurs. These three events are mutually exclusive. Let n_1 , n_2 , n_3 be the corresponding absolute frequencies in the sample. Then $\tilde{f}(A) = (n_1 + n_2)/n$, $\tilde{f}(B) = (n_1 + n_3)/n$, $\tilde{f}(A \cap B) = n_1/n$, $\tilde{f}(A \cup B) = (n_1 + n_2 + n_3)/n$. From this, (1) follows.

SECTION 18.5

5. $P(E_B | E_A) = P(E_A \cap E_B)/P(E_A) = 0.02/0.10 = 20\%$

SECTION 18.6

1. The idea of proof is the same as that of Theorem 1, but instead of filling n places we now have to fill only k places. If repetitions are permitted, we have n elements for filling each of the k places.

3. 23.5, 0.5, 2%, 39 902, 400, 1%

5. $\binom{100}{10} = 17,310,309,456,440$, cf (4a)

7. 1/1260, because there are $9!/2!3!4! = 1260$ different permutations, cf Theorem 2

SECTION 18.7

1. $f(3) = \frac{1}{81} \frac{1}{6}$, $f(4) = \frac{3}{81} \frac{1}{6}$, $f(5) = \frac{9}{81} \frac{1}{6}$, etc

3. $\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, 0, \frac{2}{6}, \frac{3}{6}, \frac{1}{6}, \frac{4}{6}$

5. $f(0) = f(4) = \frac{1}{16}$, $f(1) = f(3) = \frac{1}{8}$, $f(2) = \frac{4}{16}$

7. $f(0) = \frac{1}{8}$, $f(1) = \frac{9}{16}$, $f(2) = \frac{9}{16}$

9. $c = \alpha$

SECTION 18.8

$$\begin{aligned}1. \mu &= \int_{-\infty}^c tf(t) dt + \int_c^{\infty} tf(t) dt = - \int_{\infty}^0 (c-x)f(c-x) dx + \int_0^{\infty} (c+x)f(c+x) dx \\ &= 2c \int_0^{\infty} f(c+x) dx = c.\end{aligned}$$

For a discrete distribution the proof is similar.

3. Set $c_1 = 1/\sigma$ and $c_2 = -\mu/\sigma$ in (5) and (6).

7. $E(X - \mu) = E(X) - \mu E(1) = \mu - \mu = 0$. Analogue: The first moment with respect to the center of gravity is zero.

$$\begin{aligned} 9. \sigma^2 &= E[(X - \mu)^2] = E(X[X - 1] + X - 2\mu X + \mu^2) \\ &= E(X[X - 1]) + E(X) - 2\mu E(X) + \mu^2 \end{aligned}$$

11. Set $x - c = t$, write τ instead of t , set $\tau = -t$, and use $\mu = c, f(c - t) = f(c + t)$. Then

$$\begin{aligned} E[(X - \mu)^3] &= \int_{-\infty}^{\infty} t^3 f(c + t) dt = \int_{-\infty}^0 \tau^3 f(c + \tau) d\tau + \int_0^{\infty} t^3 f(c + t) dt \\ &= \int_{-\infty}^0 (-t)^3 f(c - t)(-dt) + \int_0^{\infty} t^3 f(c + t) dt = 0 \end{aligned}$$

13. $G^{(k)}(t) = \sum x_j^k e^{tx_j} f(x_j), G^{(k)}(0) = \sum x_j^k f(x_j) = E(X^k)$

$$G^{(k)}(t) = \int_{-\infty}^{\infty} x^k e^{tx} f(x) dx, G^{(k)}(0) = \int_{-\infty}^{\infty} x^k f(x) dx = E(X^k)$$

SECTION 18.9

1. $f(x) = \binom{5}{x} 0.5^5, f(0) = 0.03125, f(1) = 0.15625, 1 - f(0) = 0.96875, 1 - f(5) = 0.96875$

5. $f(x) = \binom{5}{x} \binom{15}{3-x} / \binom{20}{3}; f(0) = \frac{91}{240}, f(1) = \frac{35}{48}, f(2) = \frac{5}{8}, f(3) = \frac{1}{14}$

7. The probability of not hitting the target is $0.9^{10} \approx 35\%$. Ans. $1 - 0.9^{10} \approx 65\%$.

9. $f(x_1, x_2, x_3) = \frac{100!}{x_1! x_2! x_3!} (0.03)^{x_1} (0.04)^{x_2} (0.93)^{x_3}$

13. $G''(t) = n(n-1)(pe^t + q)^{n-2}(pe^t)^2 + n(pe^t + q)^{n-1}pe^t, E(X^2) = G''(0) = n(n-1)p^2 + np$, cf. Prob. 13, Sec. 18.8; $\sigma^2 = E(X^2) - \mu^2 = n(n-1)p^2 + np - n^2p^2 = npq$, cf. Prob. 8, Sec. 18.8.

15. $G''(t) = \mu e^t [G(t) + 2G'(t) + G''(t)], E[(X - \mu)^3] = \mu$, cf. Prob. 10, Sec. 18.8, $\gamma = E[(X - \mu)^3]/\sigma^3 = 1/\sqrt{\mu}$, cf. Prob. 11, Sec. 18.8.

SECTION 18.10

1. 4.8%, 79.7%, 90.5%

3. 68%, 98.8%

5. About 160

7. $P = \sum_{x=0}^{10} \binom{1000}{x} 0.01^x 0.99^{1000-x} \approx \Phi\left(\frac{10 - 10 + 0.5}{\sqrt{9.9}}\right) - \Phi\left(\frac{0 - 10 - 0.5}{\sqrt{9.9}}\right) = 0.564$ (exact 0.583)

9. $\Phi\left(\frac{4040 - 2020 + 0.5}{\sqrt{1010}}\right) - \Phi\left(\frac{2048 - 2020 - 0.5}{\sqrt{1010}}\right) = 19.3\%$

13. $P\left(\left|\frac{X}{n} - p\right| < \epsilon\right) = P([p - \epsilon]n < X < [p + \epsilon]n)$. Apply (13) with $a = (p - \epsilon)n$,

$b = (p + \epsilon)n$; then $\beta = (\epsilon n + 0.5)/\sqrt{n p q}$, $\alpha = -\beta$ and $\alpha \rightarrow -\infty$, $\beta \rightarrow \infty$ as $n \rightarrow \infty$. Hence the above probability approaches $\Phi(\infty) - \Phi(-\infty) = 1 - 0 = 1$.

SECTION 18.11

1. (X, Y) takes a value in A, B, C , or D (cf. the figure) with probability $F(b_1, b_2)$, a value in A or C with probability $F(a_1, b_2)$, a value in C or D with probability $F(b_1, a_2)$, a value in C with probability $F(a_1, a_2)$, hence a value in B with probability given by the right-hand side of (2).

$Y = b_2$		
	A	B
$Y = a_2$		
	C	D
	$X = a_1$	
	$X = b_1$	

Problem 1.

3. X_j = number of successes in the j -th trial has the mean p . Hence $X = X_1 + \dots + X_n$ number of successes in n trials has the mean $\mu = np$.

7. $f_1(x) = 1/(\beta_1 - \alpha_1)$ when $\alpha_1 < x < \beta_1$, $f_2(y) = 1/(\beta_2 - \alpha_2)$ when $\alpha_2 < y < \beta_2$

SECTION 18.12

1. The items 38, 69, 02, 49, 23, 52, 73, 29, 09, 05

SECTION 18.14

1. $\text{CONF}\{12.2 \leq \mu \leq 16.6\}$

$$3. n \approx 15\,300, n \approx 25\,000$$

7. $-D_2$ is normal with mean $-\mu_2$ and variance σ_2^2 , cf. Theorem 1, Sec. 18.10. Hence $D = D_1 - D_2$ is normal with mean $\mu = \mu_1 - \mu_2 = 0.01$ and variance $\sigma^2 = \sigma_1^2 + \sigma_2^2 = 0.0002$. A bolt fits a hole if and only if D assumes a positive value. Ans.

$$P(D > 0) = P(0 < D < \infty) = \Phi(\infty) - \Phi\left(\frac{0 - 0.01}{\sqrt{0.0002}}\right) \approx 76\%$$

9. $\text{CONF}\{0.726 \leq \mu \leq 0.751\}$

11. $c/\sqrt{n} = 0.715, z = 4.37, s = 0.157, \text{CONF}\{4.25 \leq \mu \leq 4.49\}$

13. $n = 41$

15. $\text{CONF}\{0.05 \leq \sigma^2 \leq 10\}$

17. $(n-1)s^2 = 0.01301, \text{CONF}\{0.0005 \leq \sigma^2 \leq 0.0076\}$

SECTION 18.15

1. Hypothesis $p = 0.5$, alternative $p > 0.5$ (p = probability of "head" in a single trial). If the hypothesis is true, $X = \text{Number of heads in } 4040 \text{ tosses}$ is approximately normal with $\mu = np = 2020$ and $\sigma^2 = npq = 1010$, cf. Sec. 18.10. Obtain c from $P(X \leq c) = \Phi([c - 2020]/\sqrt{1010}) = 0.95, c = 2072 > 2048$. The hypothesis is accepted.

3. $\alpha = 5\%, t = \sqrt{10}(0.811 - 0.800)/0.00735 \approx 4.7 > c = 1.83$. Hypothesis rejected.

5. $\alpha = 5\%$, (a) alternative $\mu < 1.25$, $t = \sqrt{10}(1.193 - 1.250)/0.038 \approx -4.74 < c = -1.83$. Hypothesis rejected. (b) Alternative $\mu > 1.15$, $t = \sqrt{10}(1.193 - 1.150)/0.038 \approx 3.58 > c = 1.83$. Hypothesis rejected.

9. Method B: $\bar{x} = 260, s_1^2 = 796.4$. Method A: $\bar{y} = 236, s_2^2 = 607.1, t_0 = 1.9 > c = 1.75 (\alpha = 5\%)$. We reject the hypothesis and assume that Method B is preferable.

SECTION 18.16

1. $c_1 = 1 - 2.58 \cdot 0.02/2 = 0.974$, $c_2 = 1.026$, $w_1 = 0.980$, $w_2 = 1.020$
 3. c_1 and c_2 will correspond to a significance level α which, in general, is not equal to 1%.
 5. $\mu_0 = 20.50$, $c_1 = 4.05 \cdot 5 = 20.25$, $c_2 = 20.75$

SECTION 18.17

1. $(1 - \theta)^2$
 5. $P(A; \theta) = \binom{N - N\theta}{5} / \binom{N}{5} \approx \theta^0(1 - \theta)^{5-0} = (1 - \theta)^5$, $AQ(\theta) \approx \theta(1 - \theta)^5$
 7. $\theta_1 \approx 0.01$, $\theta \approx 0.37$

SECTION 18.18

1. $K = 2$, $n = 355 + 123 = 478$, $e_1 = 478 \cdot \frac{3}{4} = 358.5$, $e_2 = 478 \cdot \frac{1}{4} = 119.5$,
 $\chi_0^2 = \frac{(355.0 - 358.5)^2}{358.5} + \frac{(123.0 - 119.5)^2}{119.5} = 0.137$. We choose $\alpha = 5\%$. Then $c = 3.84$
 (1 degree of freedom). Hence $\chi_0^2 < c$, that is, we may assume that the deviation of the ratio
 in the sample from the theoretical ratio 3:1 is due to random fluctuation.
 3. $a = 32$ because $\chi_0^2 = \frac{1}{25}[(a - 25)^2 + (50 - a - 25)^2] > c = 3.84$ when $a \geq 32$ (or
 $a \leq 18$).
 5. $\bar{x} = 99.4$, $\tilde{s} = 15.8$ (cf. p. 789), $K = 5$ (endpoints $-\infty$, 85, 95, 105, 115, ∞),
 $\chi_0^2 = 0.7 < c = 5.99$ ($\alpha = 5\%$). Hypothesis accepted.

SECTION 18.19

1. $\alpha = 5\%$, $P(X \geq 10) = \sum_{x=10}^{13} \binom{13}{x} 0.5^{13} = 4.6\% < \alpha$. Hypothesis rejected.
 3. Hypothesis $\mu = 0$. Alternative $\mu > 0$, $\bar{x} = 1.58$, $t = \sqrt{10} \cdot 1.58/1.23 = 4.06 > c = 1.83$ ($\alpha = 5\%$). Hypothesis rejected.
 7. 2 inversions, $n = 8$, $P(I \leq 2) \approx 0.1\%$. We reject the hypothesis and assume that there
 is negative trend.

SECTION 18.20

1. $y - 4.23 = -0.108(x - 40.7)$ 3. $y - 14.0 = 0.84(x - 29.5)$
 5. $\gamma = 95\%$, $c = 2.37$, $l = 2.37 \sqrt{47}/\sqrt{0.296 \cdot 7} = 11.3$, $b = 12.1$,
 $\text{CONF}\{0.8 \leq \beta \leq 23.4\}$

INDEX

A

Absolute
 class frequency 737
 convergence 622, 634
 error 25
 frequency 734
 value 22, 535
Acceleration 275
Accept 799
Acceptable
 lot 813
 quality level 813
Acceptance sampling 811
Addition
 of complex numbers 19
 of matrices 358
 of means 782
 of series 629
 of variances 783
 of vectors 240
Addition rule 750
Allowable number of defectives 811
Alternating current 71, 141
Alternative 801
 hypothesis 801
Ampère 69
Amplification 139
Amplitude 136, 457
Analytic at infinity 669
Analytic continuation 643, 650
Analytic function
 definition 163, 542
 derivatives 607

Analytic function (*Cont.*)
 line integrals 588
 mapping 564
 residues 675
 series representation 648, 663
 singularities 650, 672
 zeros 671
Analyzer, harmonic 464
Angle
 between curves 75, 323
 between vectors 242
Angular speed 275
Annulus 537
AOQ 815
AOQL 815
Approximate integration 31
Approximate solution
 of differential equations 47, 86, 149
 of eigenvalue problems 426
 of equations 26
Approximation, trigonometric 461
AQL 813
Arc 267
Arc length 269
Arc tan 23, 559
Area 315, 323
Argand diagram 18
Argument 22, 535
Asymptotic expansion 720
Asymptotically
 equal 775
 normal 797
Atmospheric pressure 46
Augmented matrix 392

Auxiliary equation 98, 116
Average
 (see Mean value)
Average outgoing quality 815
Axioms of probability 748

B

Bar chart 737
Beam 497
Beats 137
Bell-shaped curve 509, 717
Bernoulli 68
 distribution 767, 775
 equation 68
 law of large numbers 776
 numbers 654
Bessel equation 179, 483, 522
Bessel functions
 of the first kind 181
 Hankel 190
 modified 191
 Neumann 189
 orthogonality of 483
 of the second kind 189
 tables 182, 190, 524
 of the third kind 190
 zeros 182, 524
Bessel inequality 462, 475
Beta function 714
Bilinear form 413
Binomial
 coefficient 755
 distribution 767, 775, 828
 series 653
Binormal 272
Bolzano-Weierstrass theorem 614
Borda 53
Boundary
 condition 477, 487
 point 538
 value problem 477, 527, 567
Bounded
 function 84
 region 307
 sequence 614
 set 538
Bounds for eigenvalues 422
Bound vector 237
Branch point 584
Buffon 749

Cantilever beam 497
Cantor-Dedekind axiom 617

Capacitance 70
Capacitor 70
Cardano 18
Cardioid 270
Cartesian coordinates 18, 237
Catenary 270
Cauchy 87
 convergence principle 615, 622
 equation 116, 128, 177
 -Hadamard formula 636
 inequality 609
 integral formula 605
 integral theorem 596
 principal value 681, 684
 product 638
 -Riemann equations 78, 544
Center
 of gravity 309, 332
 of power series 156, 634
Central
 limit theorem 797
 moments 765
Centrifugal force 275
Centripetal acceleration 275
Chain rule 279
Change of variables 310
Characteristic
 determinant 409
 equation 98, 127, 409
 function 477, 493, 515
 polynomial 409
 value 406, 477, 515
 vector 406
Chebychev
 (see Tchebichef)
Chi-square
 distribution 795, 835
 test 818
Circle of convergence 635
Circuit 69, 140, 220, 458
Circular
 disk 537
 helix 267, 270, 273
 membrane 521
 ring 537
Circulation 346, 694, 700
Class 736
 frequency 737
 interval 736
 midpoint 736
Closed
 disk 538
 interval 617
 point set 538

- Closed (Cont.)**
 region 307
 set of orthogonal functions 476
- Coding** 743
- Coefficients**
 of a differential equation 94
 of a power series 156, 634
 of a system of equations 382
- Cofactor** 15, 369
- Coin tossing** 749
- Collatz** 424
- Collinear** 256
- Collinearity transformation** 420
- Column**
 of a determinant 369
 matrix 357
 of a matrix 357
 vector 357
- Combination** 753
- Comparison test** 623
- Comparisons, paired** 806
- Complement** 746, 750
- Complementary**
 error function 717
 Fresnel integrals 719
 sine integral 719
- Complementation rule** 750
- Complete** 476
- Complex**
 exponential function 550, 577, 650
 Fourier integral 472
 Fourier series 450
 function 539
 hyperbolic functions 554, 579
 impedance 145
 integral 589
 line integral 589
 logarithm 556, 578, 651
 number 18, 535
 number sphere 668
 plane 18
 plane, extended 570
 potential 581, 691, 696
 sequence 611
 series 618
 trigonometric functions 553, 578, 651
 variable 539
- Component** 237, 243
- Compressible fluid** 294
- Conditional probability** 751
- Conditionally convergent** 622
- Conduction of heat** 338, 501
- Confidence**
 interval 791
- Confidence (Cont.)**
 level 791
 limits 791
- Conformal mapping** 564
- Conjugate:**
 complex numbers 20
 harmonic functions 546
- Connected set** 279, 538
- Conservative field** 287, 350
- Consumer's risk** 813
- Continuation, analytic** 643, 650
- Continuity**
 of a complex function 540
 equation 295
 of a function of two variables 278
 piecewise 195
 of a vector function 263
- Continuous distribution** 759
- Continuously differentiable** 269
- Contour integral** 601
- Control**
 chart 809
 limit 809
 mechanism 230
- Convergence**
 absolute 622, 634
 circle of 635
 conditional 622
 interval of 160, 636
 in the mean 476
 principle 615, 622
 radius 161, 635
 of a sequence 263, 611
 of a series 160, 619
 tests for 620, 623
 uniform 657
- Cooling** 52
- Coordinate transformations** 288
- Coordinates**
 Cartesian 18, 237
 cylindrical 527
 polar 22
 spherical 527
- Coplanar** 256
- Coriolis acceleration** 276
- Correlation analysis** 822
- Cosecant** 4, 553
- Cosine**
 of a complex variable 553, 579, 651
 hyperbolic 5, 554
 integral 719
 of a real variable 2
- Cotangent** 4, 553

Index

- Coulomb 69
law 287
- Covariance 823
- Cramer's rule 11, 13, 384
- Critical
damping 114
point 566
- Cross product 247
- Cumulative
absolute frequency 735
distribution function 758
relative frequency 735
- Curl 296, 343, 346
- Curvature 272
- Curve
arc length of 269
orientation 269
piecewise smooth 301
plane 266
rectifiable 268
simple 267
smooth 300, 564
twisted 266
- Cylinder, flow around 699
- Cylindrical coordinates 527
- D**
- D'Alembert 500
- Damping
constant 111
critical 114
force 111
- Decreasing sequence 616
- Decrement 115
- Dedekind 617
- Defective item 732
- Definite complex integral 588
- Deformation of path 600
- Degrees of freedom 794, 795, 807
- Delta, Kronecker 290
- De Moivre
formula 23
limit theorem 775
- Density 760, 778
- Dependent 100, 124, 255, 780
- Derivative
of a complex function 541, 607
directional 282
left-hand 438
partial 8
right-hand 438
* of a vector function 264
- Desargues 18
- Determinant
characteristic 409
definition 10, 370
of a matrix 379
of a system of equations 382
- Development around different centers 642
- Diagonal matrix 360
- Differentiable complex function 541
- Differential
exact 348
form 348
total 348
- Differential equation
Bernoulli 68
Bessel 179, 483, 522
Cauchy 116, 128, 177
Cauchy-Reimann 78, 544
of circuits 69, 141
with constant coefficients 97, 126
elliptic 501
Euler 116, 177
exact 59
homogeneous 57, 63, 94, 123
hyperbolic 501
hypergeometric 178
Laguerre 482
Laplace 287, 339, 526, 546
Legendre 164, 530
linear 63, 94, 123
nonhomogeneous 63, 94, 123
nonlinear 94, 123
numerical methods for 86, 149
order of 40, 486
ordinary 39
parabolic 501
partial 486
Poisson 487
separable 49
series solution of 155
Sturm-Liouville 477
of vibrating beam 497
of vibrating mass 110, 134, 411, 455
of vibrating membrane 510
of vibrating string 490
Weber 482
- Differential
form 348
geometry 266
- Differentiation
analytic functions 607
asymptotic series 728
complex functions 541
Laplace transforms 198, 213
order of 9

- Differentiation (Cont.)**
- power series 643
 - series 661
 - vector functions 264
- Directed line segment** 236
- Directional derivative** 282
- Direction field** 47
- Dirichlet** 468
 - discontinuous factor 468
 - problem 705
- Discharge of a source** 698
- Discrete random variable** 757, 777
- Disjoint events** 746
- Disk** 537
- Distribution** 758
 - Bernoulli 767, 775
 - binomial 767, 775, 828
 - chi-square 795, 835
 - continuous 759
 - discrete 757, 777
 - Fisher's F- 807, 836
 - function 736, 737, 758
 - Gauss 771
 - hypergeometric 768
 - marginal 779
 - multinomial 770
 - normal 771, 789, 791, 803, 830
 - Poisson 768, 829
 - rectangular 763
 - Student's *t*- 794; 834
 - two-dimensional 777
 - uniform 763
- Divergence**
 - theorem of Gauss 333
 - of vector fields 292, 337
- Divergent** 160, 612, 619
- Division of complex numbers** 19
- Domain** 279, 538
- Dot**
 - frequency diagram 737
 - product 242
- Double**
 - Fourier series 516
 - integral 307, 313
 - point 267
 - sampling plan 816
- Driving force** 134
- Dyne** 109
- E**
- Eigenfunction**
 - expansion 480
 - of Sturm-Liouville problem 477
- Eigenfunction (Cont.)**
 - of vibrating membrane 515
 - of vibrating string 493
- Eigenvalue**
 - of matrix 406
 - of Sturm-Liouville problem 477
 - of vibrating membrane 515
 - of vibrating string 493
- Eigenvector** 406
- Electrical network** 230, 385, 404
- Electric circuit** 69, 140, 220, 458
- Electromechanical analogies** 142
- Electromotive force** 69
- Electrostatic**
 - field 689
 - potential 689
- Element**
 - of area 323
 - of a determinant 369
 - of a matrix 357
 - of a sample space 745
- Elimination method** 399
- Elimination of first derivative** 184
- Ellipse** 267, 274
- Ellipsoid** 320
- Elliptic**
 - differential equation 501
 - paraboloid 320
- Empty event** 746
- Entire function** 673
- Equality**
 - of complex numbers 19
 - of matrices 357
 - of vectors 236, 239
- Equation**
 - characteristic 98, 127, 409
 - indicial 169
 - quadratic 26
 - solution of 26
- Equipotential**
 - lines 690, 696
 - surfaces 690
- Error** 25
 - of Type I and II 803
- Error function** 509, 716, 723
- Escape from the earth** 52
- Essential singularity** 673
- Estimation of parameters** 785
- Euler** 87
 - Beta function 714
 - Cauchy method 87, 88
 - constant 188
 - equation 116, 177
 - formula 551, 554

- Euler (Cont.)**
- formulas for Fourier coefficients
 - 435, 445
 - numbers 655
- Evaporation** 56
- Even function** 2, 440
- Event** 745
 - disjoint 746
 - empty 746
 - impossible 746
 - independent 751
 - mutually exclusive 746
- Exact**
- differential 59, 348
 - differential equation 59
- Existence theorem**
- differential equations 84, 118, 124
 - Fourier integrals 467
 - Fourier series 438
 - Laplace transforms 196
- Expectation** 765, 782
- Experiment** 744
- Exponential**
- decay 56
 - function, complex 550, 577, 650
 - function, real 1, 650
 - growth 56
 - integral 730
- Extended**
- complex plane 570
 - power series method 168
- F**
- F-distribution** 807, 836
- Factor, integrating** 61
- Factorial function** 711, 754
- Failure** 767
- Fair die** 750
- Falling body** 46
- Family of curves** 74
- Faraday** 69
- Field**
- conservative 287, 350
 - of force 261
 - gravitational 261, 286, 526
 - irrotational 695
 - scalar 260
 - vector 261
 - velocity 261, 693
- Finite complex plane** 570
- First**
- fundamental form 322
 - Green's formula 340
 - shifting theorem 195
- Fisher's F-distribution** 807, 836
- Fixed point** 569
- Fluid flow** 693
- Flux** 294
- Foot** 109
- Forced oscillations** 134, 210, 225, 455, 497
- Form**
- bilinear 413
 - fundamental 322
 - Hermitian 414
 - quadratic 413
 - skew-Hermitian 415
- Fourier** 431
 - Bessel series 485
 - coefficients 436, 445, 453
 - coefficients, complex 450
 - constants 474
 - cosine series 441
 - double series 516
 - half-range expansions 448
 - integral 467, 507, 683
 - integral, complex 472
 - series 436, 494
 - series, complex 450
 - series, generalized 474
 - sine series 441
 - transform 472
- Four-terminal network** 404
- Fractional linear transformation** 570
- Fraction defective** 812
- Free oscillations** 108, 210
- Frenet formulas** 273
- Frequency** 734
 - absolute 734
 - cumulative 735
 - cumulative relative 735
 - distribution 736
 - function of a sample 736
 - histogram 737
 - of oscillation 111
 - polygon 737
 - relative 735
- Fresnel integrals** 717
- Frobenius**
- method of 169
 - theorem of 424
- Full coding** 743
- Full wave rectification** 227
- Function**
- analytic 163
 - Bessel
 - (see Bessel functions)
 - beta 714
 - bounded 84

Function (Cont.)

- characteristic 477, 493, 515
- complex 539
- conjugate harmonic 546
- entire 673
- error 509, 716, 723
- even 2, 440
- exponential 1, 550, 577, 650
- factorial 711
- gamma 180, 710
- Hankel 190
- harmonic 526, 546, 567, 703
- hyperbolic 5, 554, 573, 651
- inverse hyperbolic 559
- inverse trigonometric 559
- Legendre 164
- logarithmic 1, 556, 578, 651
- meromorphic 673
- multi-valued 539
- Neumann 189
- odd 2, 441
- orthogonal 473, 475, 483
- orthonormal 473, 475
- periodic 223, 431
- probability 757
- rational 548
- scalar 260
- single-valued 539
- staircase 217, 224
- step 216
- trigonometric 2, 553, 578, 651
- unit step 216
- vector 261

Fundamental

- form 322
- mode 493
- system 101, 125, 396
- theorem of algebra 610
- triad 241

Funnel 53**G**

- Gamma function** 180, 710
- Gauss 18**
 - algorithm 399
 - distribution 771
 - divergence theorem 333
 - elimination method 399
 - hypergeometric equation 178
- General**
 - exponential function 558
 - powers 558
 - solution 42, 100, 125

Generalized Fourier series 474**Geometric series 161, 623, 650****Gershgorin 422****Gibbs phenomenon 469****Goodness of fit 817****Goursat 596****Gradient 283****Gravitational field 261, 286, 526****Green****formulas 340****theorem 313, 340, 346****Grouping 630, 736****Guarding figures 25****H****Hadamard's formula 636****Hankel functions 190****Half-life time 55****Half-range Fourier series 448****Half-wave rectifier 223, 446****Harmonic****analyzer 464****conjugate 546****function 342, 526, 546, 567, 703****oscillation 111****series 619****Heat****equation 339, 502****flow 338, 501, 580****Helix 267, 270, 273****Helmholtz equation 529****Henry 69****Hermite polynomials 481****Hermitian****form 414****matrix 414****Hesse's normal form 245****High frequency line equations 533****Histogram 737****Holomorphic 542****Homogeneous****differential equation 57, 63, 94, 123****system of equations 12, 382****Höne (Wronsky) 120****Hooke's law 110****Hyperbolic****differential equation 501****functions, complex 554, 479, 651****functions, real 5****paraboloid 320****Hyperboloid 320****Hypergeometric****differential equation 178**

Hypergeometric (Cont.)
 distribution 768
 functions 178
 series 178
Hypocycloid 270
Hypothesis 799

I

Identity
 of Lagrange 259
 of power series 641
 transformation 568
Image 561
Imaginary
 axis 18
 part, 18, 535
 unit 18
Impedance 142, 145
Impossible event 746
Improper integral 681, 684
Incomplete gamma functions 714
Incompressible 295, 695
Increasing sequence 616
Indefinite
 integral 588
 integration 604
Independence
 linear 100, 124, 255
 of path 348, 599
Independent
 events 751
 random variables 780
Indicial equation 169
Inductance 70
Inductor 70
Inequality
 Bessel 462, 475
 Cauchy 609
 Schur 424
 triangle 536
Infinite
 population 769
 sequence 611
 series 618
Infinity
 analytic at 669
 point at 668
Initial
 condition 51, 124, 487
 value problem 51, 104, 117
Input 66, 134
Insertion of parentheses 630
Instrumental methods 32, 464

Integral
 contour 601
 cosine 719
 definite 588
 double 307, 313
 exponential 730
 Fourier 467, 507, 683
 Fresnel 717
 improper 681, 684
 indefinite 588
 line 301, 589
 logarithmic 730
 sine 468, 719
 surface 327
 triple 333
Integral theorems
 complex 596, 605
 real 313, 333, 340, 342
Integrating factor 61
Integration
 asymptotic series 727
 complex functions 589
 Laplace transforms 199, 213
 power series 643
 series 660
Interquartile range 744
Intersection of events 746
Interval
 closed 617
 of convergence 160, 636
 estimate 790
 open 617
Inverse
 hyperbolic functions 559
 Laplace transform 193
 of a matrix 402
 trigonometric functions 559
Inversion 569
Irrotational 695
Isoclone 47
Isogonal trajectories 78
Isolated
 point 672
 singularity 673
Iteration
 for eigenvalues 426
 for equations 29
 Picard 79

Jacobian 311, 566
Joukowski airfoil 586
Jump 451

K

Kirchhoff's laws 70
 Kronecker delta 290
 Kutta 89

L

Lagrange 67
 identity of 259
 Laguerre polynomials 482
 Lambert's law 56
 Laplace 193
 equation 287, 339, 526, 546
 integrals 470
 limit theorem 775
 operator 287
 transform 193
 Laplacian 287, 316, 519, 528
 Laurent series 663
 Law
 of cooling 52
 of gravitation 261
 of large numbers 776
 of mass action 56
 of the mean 280, 307, 337

LCL 809

Least squares 822

Left-hand
 derivative 438
 limit 438

Left-handed
 coordinates 249
 triple of vectors 249

Legendre 164
 differential equation 164, 530
 functions 164
 polynomials 166, 476, 480, 481, 531

Leibniz 27

test 620

Length

of a curve 268
 of a vector 236

Level

curve 561
 surface 262, 284

Likelihood function 788

Limit

of complex function 540
 left-hand 438
 point 613, 672
 right-hand 438
 of sequence 263, 611
 theorem 775
 of vector function 263

Line integral

complex 589
 real 301, 313, 342

Line of force 691

Lineal element 47

Linear

combination 255
 dependence 100, 124, 255
 differential equation 63, 94, 123
 element 269
 independence 100, 124, 255
 mapping 570
 transformation 570

Liouville 477

theorem of 609

Lipschitz condition 85

Logarithm 1, 556, 578, 651

Logarithmic

decrement 115
 integral 730

Lot tolerance per cent defective 813

Lower

control limit 809
 quartile 744

LTPD 813

M

M-test 661

Maclaurin 550
 series 648

Magnitude of a vector 236

Mapping 560

conformal 564

Marginal distribution 779

Mass action 56

Mathematical expectation 765, 782

Matrix 357

addition 358
 augmented 392
 column 357
 diagonal 360

Hermitian 414

multiplication 363

nonsingular 380

normal 424

null 361

orthogonal 417

real 359

row 357

scalar 367

singular 380

skew-Hermitian 415

skew-symmetric 359, 418

- Matrix (Cont.)**
 square 359,
 symmetric 359, 418
 transpose of 359, 367, 380
 triangular 360
 unit 360
 unitary 417
 zero 361
- Maximum**
 likelihood method 788
 modulus theorem 703
 principle 704
- Mean**
 of distribution 762
 of sample 740
- Mean value**
 of analytic function 703
 of distribution 762
 of function 694
 of harmonic function 704
 of sample 740
 theorem 280, 307, 337
- Median** 744
- Membrane** 510
- Meromorphic function** 673
- Method**
 Euler-Cauchy 87, 88
 extended power series 168
 of false position 28
 Frobenius 169
 of iteration 29, 79, 426
 of least squares 822
 of moments 785
 of numerical integration 31
 power series 155
 Runge-Kutta 89, 151
 of undetermined coefficients 131, 654
- Middle quartile** 744
- Minor** 15, 369
- Mixed triple product** 254
- Möbius strip** 328
- Modal matrix** 420
- Mode** 493, 523
- Modified Bessel functions** 191
- Modular surface** 555
- Modulus** 22
- Moivre's formula** 23
- Moment**
 generating function 765
 of a distribution 765
 of a force 251
 of inertia 309, 312, 328
 of a sample 785
 vector 252
- Monotone**
 decreasing 616
 increasing 616
 sequence 617
- Morera's theorem** 609
- Moving trihedron**
(see Trihedron)
- Multinomial distribution** 770
- Multiple point** 267
- Multiplication**
 asymptotic series 726
 complex numbers 19
 matrices 363
 means 782
 power series 638
 vectors 240, 242, 247, 254
- Multiplication rule** 751
- Multiplicity** 410
- Multiply connected** 595
- Multi-valued function** 539
- Mutual inductance** 230
- Mutually exclusive events** 746
- N**
- Nabla** 283
- Natural**
 frequency 135
 logarithm 1, 556, 578, 651
- Neighborhood** 263, 278, 537
- Networks** 230, 385, 404
- Neumann's function** 189
- Newton** 27
 law of cooling 52
 law of gravitation 261
 method of 27
 second law 110
- Nodal line** 515
- Node** 493
- Nondefective item** 732
- Nonhomogeneous**
 differential equation 63, 94, 123, 128
 systems of equations 12, 382
- Nonlinear differential equation** 94, 123
- Nonorientable surface** 328
- Nonparametric test** 819
- Nonsingular matrix** 380
- Nontrivial solution** 386
- Norm** 473, 475
- Normal**
 asymptotically 797
 to a curve 272
 distribution 771, 789, 791, 803, 830
 matrix 424
 mode 493, 523

- N**ormal (*Cont.*)
 probability paper 786
 random variable 771
 to a plane 245
 to a surface 284, 321
 vector 244
- N**ull
 matrix 361
 vector 238
- N**umber sphere 669
- N**umerical methods
 differential equations 86, 149
 eigenvalues 422
 equations 26
 Fourier coefficients 458
 integration 31
 linear equations 399
- N**yström 151
- O**
- O**bervation 744
Observed value 756
OC curve 804, 812
Odd function 2, 441
Ohm 69
One-dimensional
 heat equation 502
 wave equation 490
- O**ne-sided
 alternative 801
 test 802
- O**ne-to-one mapping 566
- O**pen
 disk 537
 interval 617
 set 538
- O**perating characteristic 804, 812
- O**perations on series 629
- O**rder
 determinant 370
 differential equation 40
 differentiation 9
 eigenvalue 409
 square matrix 359
- O**rdinary differential equation 39
(see also Differential equation)
- O**riental surface 328
- O**rientation of a curve 269, 564
- O**rthogonal
 curves 75
 functions 473, 475, 483
 matrix 417
 net 75
- O**rthogonal (*Cont.*)
 trajectories 75
 triad 262
 vectors 241
- O**rthonormal 473, 475
- O**scillations
 of beam 497
 in circuits 71, 141
 damped 111, 136
 forced 134, 210, 225, 455
 free 108, 210
 harmonic 111
 of mass on a spring 108, 134, 411, 421, 455
 of membrane 510
 of string 488, 497
 undamped 110, 135
- O**utcome 744
Output 66, 134
Overdamping 112
Overtone 493
- P**aired comparisons 806
Pappus, theorem of 326
Parabola, semicubical 270
Parabolic differential equation 501
Paraboloid 320
Parallel flow 697
Parallelepiped 254
Parallelogram law 19, 240
Parameter 785
Parameters, variation of 67
Parametric representation 266, 319
Parseval's formula 476
Partial
 derivative 8
 differential equation 486
 fraction 204
 sum 159, 618
- P**articular solution 42, 100, 125
Path of integration 301, 589
- P**earson 749
- P**ercentile 744
- P**eriodic function 223, 431
- P**ermutation 752
- P**hase
 angle 139
 of complex number
(see Argument)
 lag 139
- P**icard
 iteration method 79
 theorem 674

- Piecewise 1
 continuous 195
 smooth 301, 319, 590
- Plane, complex 18, 570
- Plane curve 266
- Planimeter 32
- Point
 estimate 790
 at infinity 570
 multiple 267
 set 538
 source 698
- Poisson 705
 distribution 768, 829
 equation 487
 integral formula 706
- Polar
 coordinates 22
 form 22
 moment of inertia 309
- Pole 672
- Polynomial 548
 Chebychev 481
 Hermite 481
 Laguerre 482
 Legendre 166, 476, 480, 481
 Tchebichef 481
 trigonometric 461
- Population 733, 784
- Position vector 239
- Positive definite 416
- Possible value 759
- Postmultiplication 366
- Potential
 complex 581, 691, 696
 real 286, 526
 theory 526, 689
- Pound 109
- Power 802
 function 802
- Powers 558
- Power series 155, 159, 633
- Power series method 155
- Premultiplication 366
- Primitive period 431
- Principal
 diagonal 359, 369
 normal 272
 part 672
 value 535, 549, 556, 681, 684
- Probability 749
 axioms of 749
 conditional 751
 density 760, 778
- Probability (Cont.)
 distribution 758
 function 757, 778
 paper 786
- Producer's risk 813
- Product
 of asymptotic series 726
 of complex numbers 19
 of determinants 376
 of matrices 363
 of power series 638
 of vectors 240, 242, 247, 254
- Pure imaginary number 21
- Q**
- Quadratic
 equation 26
 form 413
- Quality control 809
- Quartile 744
- Quotient of complex numbers 20
- R**
- Radiation of radium 44, 56
- Radius of convergence 161, 635
- Random
 digits 784, 832
 experiment 744
 numbers 784
 observation 744
 sample 734
 variable 756
- Range of values 539
- Rank of a matrix 380
- Rational function 548
- Ratio test 624, 627
- RC-circuit 72, 220
- Reactance 142
- Real
 axis 18
 matrix 359
 part 18, 535
- Rearrangement of series 631
- Reciprocal of a matrix 402
- Rectangular
 distribution 763
 membrane 512
 rule 33
- Rectifiable 268
- Rectification 814
- Rectifier 223, 227, 446
- Reduction of order 96, 128
- Region 538

- Regression 822
 - coefficient 823
 - line 823
- Regula falsi 28
- Regular 542
- Reject 799
- Relative
 - class frequency 737
 - error 25
 - frequency 735
- Remainder 159, 618
- Removable singularity 673
- Representation, parametric 266, 319
- Residue 675
 - Residue theorem 678
- Resistance 69
- Resonance 136, 210
- Response 66, 134
- Resultant of forces 240
- Riemann 544
 - equations 544
 - number sphere 669
 - surface 584
- Right-hand
 - derivative 438
 - limit 438
- Right-handed
 - coordinates 249
 - triple of vectors 247
- Risk
 - consumer's 813
 - producer's 813
- RL*-circuit 70, 222
- RLC*-circuit 141, 228, 458
- Rodrigues, formula of 167
- Root 548
- Root test 626, 628
- Rotation 252, 276, 297, 695
- Rounding error 25
- Row
 - of a determinant 369
 - matrix 357
 - of a matrix 357
 - vector 357
- Runge-Kutta method 89, 151

- Sample 734, 784
 - cumulative frequency function 736
 - distribution function 736
 - frequency function 736
 - mean 740
 - size 734
- Sample (*Cont.*)
 - space 745
 - value 734
 - variance 740
- Sampling
 - plan 811, 816
 - with replacement 751, 768
 - without replacement 752, 768
- Saw-tooth wave 224
- Scalar 236
 - field 260
 - function 260
 - matrix 367
 - product 242
 - triple product 254
- Schur's inequality 424
- Secant 4, 553
- Second Green's formula 340
- Second shifting theorem 220
- Sectionally continuous
 - (*see* Piecewise continuous)
- Semicubical parabola 270
- Separation of variables 49, 490
- Sequence 611
- Series 618
 - addition of 629
 - asymptotic 720
 - of Bessel functions 485
 - binomial 653
 - convergence 160, 619
 - differentiation 643, 661
 - double Fourier 516
 - of eigenfunctions 480
 - Fourier 436, 494
 - geometric 161, 623
 - harmonic 619
 - hypergeometric 178
 - infinite 618
 - integration 643, 660
 - Laurent 663
 - Maclaurin 648
 - multiplication of 638
 - of orthogonal functions 480
 - partial sums of 618
 - power 159, 633
 - remainder of 159, 618
 - Taylor 648
 - trigonometric 432
 - value of 619
- Serret-Frenet formulas
 - (*see* Frenet formulas)
- Set of points 538
- Shifting theorems 195, 220
- Sign test 820

- Significance level** 801
Significant 799
Simple
 - curve 267
 - pole 676**Simply connected** 351, 595
Simpson's rule 36
Simultaneous
 - differential equations 211, 407, 411
 - linear equations 10, 382**Sine**
 - of complex variable 553, 578, 651
 - hyperbolic 5, 554, 579
 - integral 468, 719
 - of real variable 2**Single sampling plan** 811
Single-valued function 539
Singular
 - at infinity 669
 - matrix 380
 - point 650, 672
 - solution 42**Singularity** 650, 672
Sink 338, 695, 698
Size of sample 734
Skew-Hermitian
 - form 415
 - matrix 415**Skew-symmetric matrix** 359, 418
Skewness 765
Sliding vector 237
Slug 109
Smooth
 - curve 300, 564
 - piecewise 301, 319, 590
 - surface 319**Solution**
 - of differential equation 40, 95, 124, 487
 - of equations 26
 - general 100, 125
 - particular 100, 125
 - of quadratic equation 26
 - singular 42
 - steady state 72
 - of systems of equations 382, 386
 - transient 71**Source** 338, 695, 698
Spectral matrix 420
Spectrum 406, 493
Speed 275
 - angular 275**Sphere** 319
Spherical coordinates 527
Spring
 - force 110
 - modulus 110**Square**
 - error 461
 - matrix 359
 - wave 217, 436**Stability of frequencies** 748
Stagnation point 699
Staircase function 217, 224
Standard deviation
 - of distribution 763
 - of sample 740**Standardized variable** 764
Statistic 740
Statistical regularity 748
Statistics 732
Steady state 71, 136
Steiner's theorem 332
Step function 216
Stereographic projection 668
Stirling formula 713, 754
Stochastic variable 756
Stokes's theorem 342
Straight line 244, 266
Stream function 696
Streamline 693
Strength of a source 698
String, vibrating 489
Student's *t*-distribution 794
Sturm-Liouville
 - equation 477
 - problem 477**Subevent** 747
Submarine cable equations 533
Submatrix 380
Subsidiary equation 202
Subtraction
 - of complex numbers 19
 - of matrices 358
 - of series 629
 - of vectors 241**Success** 767
Sum
 - of complex numbers 19
 - of matrices 358
 - of normal variables 795
 - of a series 619
 - of vectors 240**Superposition principle** 95
Surface 318
Surface integral 327
Symmetric
 - distribution 762

- Symmetric (Cont.)**
 matrix 359, 419
- System**
 of differential equations 211, 407, 411
 of linear equations 10, 382
- T**
- t*-distribution** 794, 834
- Tables of functions**
 arc tan 23
 Bessel functions 182, 190, 524
 cosine 6
 cosine integral 718
 error function 718
 exponential function 6
 factorial function 755
 gamma function 715
 hyperbolic functions 6
 natural logarithm 6
 sine 6
 sine integral 718
 tangent 6
 trigonometric functions 6
- Table of Laplace transforms** 232
- Tables, statistical**
 binomial distribution 828
 chi-square distribution 835
 F-distribution 836
 normal distribution 830
 Poisson distribution 829
 random digits 832
 t-distribution 834
 test for trend 840
- Tally chart** 734
- Tangent** 4, 553
 hyperbolic 5, 555
- Tangent**
 to a curve 270
 plane 284, 321
 vector 271
- Taylor**
 formula 647
 series 648
- Tchebichef polynomials** 481
- Telegraph equations** 533
- Test**
 chi-square 818
 for convergence 620, 623
 of hypothesis 799
 nonparametric 819
- Tetrahedron** 256
- Thermal conductivity** 338
- Three-sigma limits** 772
- Time constant** 73
- Torricelli's law** 53
- Torsion of a curve** 273
- Torus** 325
- Total**
 differential 59, 348
 square error 461
- Trace of a matrix** 420
- Trajectories** 75, 78
- Transformation**
 of Cartesian coordinates 288, 418
 by a complex function 560
 of integrals 313, 333, 342
 linear 570
 of vector components 289
- Transient state** 71, 136
- Translation** 236, 568
- Transmission line equations** 532
- Transpose of a matrix** 359, 367, 380
- Trapezoidal rule** 34
- Trend** 820
- Triad, fundamental** 241
- Triangle inequality** 536
- Triangular matrix** 360
- Trigonometric**
 approximation 461
 form of complex numbers 22
 functions, complex 553, 578, 651
 functions, real 2
 polynomial 461
 series 432
- Trihedron** 272
- Triple integral** 333
- Triple of vectors** 247
- Trivial solution** 384, 386
- Truncation error** 25
- Twisted curve** 266
- Two-dimensional**
 distribution 777
 random variable 777
 wave equation 512
- Two-sided**
 alternative 801
 test 802
- Type I and II errors** 802
- U**
- UCL*** 809
- Unacceptable lot** 813
- Undamped system** 110
- Underdamping** 112
- Undetermined coefficients** 131, 654
- Uniform**
 distribution 763
 convergence 657

- Union of events** 746
- Uniqueness**
 - differential equations 84, 118, 124, 340, 705
 - Laurent series 666
 - power series 641
- Unit**
 - binormal vector 272
 - circle 537
 - matrix 360
 - normal vector 245, 321
 - principal normal vector 272
 - step function 216
 - tangent vector 271
 - vector 237
- Unitary**
 - matrix 417
 - system of vectors 417
- Upper**
 - control limit 809
 - quartile 744
- V**
- Value of a series** 619
- Variable**
 - complex 539
 - random 756
 - standardized 764
 - stochastic 756
- Variance**
 - of distribution 762
 - of sample 740
- Variate** 756
- Variation of parameters** 67, 147
- Vector** 236
 - addition 240
- Vector (*Cont.*)**
 - field 261
 - function 261
 - moment 252
 - product 247
- Velocity**
 - of escape 53
 - field 261, 693
 - potential 695
 - vector 252, 275
- Venn diagram** 745
- Vibrations**
 - (*see* Oscillations)
- Volta** 69
- Voltage drop** 69
- Volume** 309
- Vortex** 698
- Vorticity** 695
- W**
- Wave equation** 490, 512
- Weber's equation** 482
- Weierstrass** 544
 - M-test 661
- Weight function** 475
- Wessel** 18
- Wheatstone bridge** 386
- Work** 244, 304
- Working origin** 742
- Wronskian** 120, 125
- Z**
- Zero**
 - of analytic function 671
 - matrix 361
 - vector 238

